REMARKS

Claims 3, 11, 12, 13, 17, 18, 19, 20, 22, 30, 31, 32, 34, 35, 36, 44, 47, 48, and 49 have been cancelled. Claims 37-40 and 42 have been amended. New claims 50-58 have been added. Claims 1, 2, 4-10, 15-16, 21, 23-29, 37-40, 42, and 50-58 are now pending in the application. No new matter has been added by amendment. Reexamination and reconsideration of the claims as amended are respectfully requested.

Claim Rejections – 35 USC § 112, second paragraph

The Examiner rejects claims 3, 11-13, 18-20, 22, 30-32, and 47-49 under 35 U.S.C. 112 second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The Applicant traverses the rejection.

The Examiner rejects claims 3 and 22. Claims 3 and 22 have been cancelled.

The Examiner rejects claims 11 and 30, and dependent claims 12-13 and 31-32. Claims 11, 12, 13, 30, 31, and 32 have been cancelled.

The Examiner rejects claims 18 and 47, and dependent claims 19-20 and 48-49. Claims 18, 19, 20, 47, 48, and 49 have been cancelled.

The Examiner states that the dependent claims cited in this rejection fail to further limit the claims from which they depend. The Examiner suggests that the claims be placed in a product –by-process format. New claims 51-58 reflect that suggestion. The Examiner also suggests that the claims should be drafted in terms of methods of making a plant by comprising transforming the exemplified plant of claim 2 or 21. New claims 55-58 reflect that suggestion.

Claim Rejections – 35 USC § 112, first paragraph

The Examiner rejects claims 3, 9-13, 15-20, 22, 28-32, 34-40, 42, 44, and 47-49 under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to

one skilled in the art that the inventors, at the time the application was filed, had possession of the claimed invention. The Applicant traverses the rejection. Claims 3, 11, 12, 13, 17, 18, 19, 20, 22, 30, 31, 32, 34, 35, 36, 44, 47, 48, and 49 have been cancelled. Claims 37, 38, 39, 40, and 42 have been amended. New claims 50-58 have been added.

The Examiner rejects claims 11, 15-16, 30, 34-35 and their dependent claims and states that the claims "are broadly drawn to any transgenic plant which contains any heterologous transgene of any sequence conferring any trait, and methods of making and using the transgenic plants." The Examiner rejects claims 3, 18-20, 22, and 47-49 and their dependent claims and states that the claims "are broadly drawn to any 'single gene conversion' plant comprising one or more traits including male sterility introgressed into the claimed variety by backcrossing or other traditional means, and methods of using these plants." The Examiner states that "no guidance has been provided for the introgression of any single trait from a multitude of non-disclosed and uncharacterized parentals into the claimed variety, wherein said introgression should result in successful expression of the desired trait but should not interfere with the expression of the remaining traits whose combination confers patentability to the instantly exemplified variety, and which introgression should not introduce unwanted linked genetic material into the exemplified cultivar which would disrupt its patentably unique genetic complement."

Claims 3, 11, 12, 13, 18, 19, 20, 22, 30, 31, 32, 47, 48, and 49 have been cancelled and new claims 51-58 have been added. The new claims are in the method and product-by-method format requested by the Examiner. The claims include the well known methods of producing backcross and transgenic conversion plants. The product-by-process claims are further limited by specified conversion or transgenic traits, which include the traits of insect resistance, herbicide resistance, disease resistance, waxy starch, and male sterility.

While the Examiner states that the claims encompass "any transgenic plant which contains any heterologous transgene of any sequence conferring any trait", the Applicant points out that it is not claiming so broadly as to claim any maize plant, regardless of source, comprising those traits. The Applicant is claiming PH0R8 or a limited set of plants derived therefrom that have obtained significant genetic contribution from PH0R8.

Applicant respectfully points out that examples of transgenes, genes, and traits that can be backcrossed into the PH0R8 are given in the application on page 20, lines 16-34, and also on page 22, line 20, through page 32, line 4. In order to expedite prosecution new claims 52 and 56 list the type of traits that may be conferred by backcross conversions and transgenes. Claim 52 also specifies that PH0R8 is used at least twice as a recurrent parent in the development of a backcross conversion plant. Breeders, by using molecular markers, may obtain up to 98% genome identity between the backcross conversion and the recurrent parent after two backcrosses. See Marker-assisted Selection in Backcross Breeding, Openshaw, S.J. et al. Marker-assisted selection in backcross breeding. In: Proceedings Symposium of the Analysis of Molecular Data, August 1994, pp. 41-43. Crop Science Society of America, Corvallis, OR (1994) included as Appendix A. Inbred PH0R8 transformed to comprise a transgene is also easily identifiable through the use of molecular markers. The transgenic version of PH0R8 would have the same molecular profile as PH0R8, with the possible exception of a marker used in the profile that is located at the site of transgene insertion. However, in this case, the plethora of other identical markers would identify the line as a transgenic variant of PH0R8.

Applicant points out to the Examiner that, at the present time, it is not typical that a transgene be incorporated into each newly developed line, such as PH0R8, by direct transformation. Rather, the more typical method used by breeders of ordinary skill in the art is to incorporate the transgene into a new line by taking an already transformed plant line and using it as a donor line to produce a backcross conversion. Thus, the well established method of

backcrossing has been used and is the most common means of introgressing the claimed traits into new material.

In the specification on page 4, lines 7-13, it states, "Backcrossing can be used to transfer a specific desirable trait from one inbred or source to an inbred that lacks that trait. This can be accomplished, for example, by first crossing a superior inbred (recurrent parent) to a donor inbred (non-recurrent parent), that carries the appropriate gene(s) for the trait in question. The progeny of this cross is then mated back to the superior recurrent parent followed by selection in the resultant progeny for the desired trait to be transferred from the non-recurrent parent." The method of backcrossing genes into an inbred maize plant is well known and well understood to one of ordinary skill in the art. The method has been successfully used since the 1950's (see pages 585-586 of Wych, 1988 included in the Information Disclosure Statement). In the specification, on page 20, lines 16-34, there is a description of how to backcross traits into PH0R8, which includes the claimed traits. Examples of how one of ordinary skill in the art can transfer a gene conferring a qualitative trait into a variety through backcrossing is demonstrated by the fact that the commercial market now distributes a multitude of products produced in this manner. Such conversion lines are easily developed without undue experimentation. Poehlman et al. (1995) on page 334, submitted in the information disclosure statement, states that, "A backcross-derived inbred line fits into the same hybrid combination as the recurrent parent inbred line and contributes the effect of the additional gene added through the backcross." Wych (1988) on page 585-86, also submitted in the information disclosure statement, discusses how the male sterility trait is routinely backcrossed into an inbred line and how this is used to produce a sterile/fertile blend of an F1 hybrid in order to reduce seed production costs. In fact, many commercial products are produced in this manner, and those of ordinary skill in the art consider the F1 hybrid produced with the male sterile (backcross conversion) inbred to be the same variety as the F1 hybrid produced with the non-backcross conversion inbred.

As a result of the repeated use of the recurrent parent, the backcross conversion has many genetic alleles in common with the recurrent parent. Thus, genetic analysis may be used as a means of identifying the backcross conversion. The F1 hybrid made with a transgenic version or a backcross conversion of PH0R8 is also identifiable by the use of genetic markers, because the hybrid would contain one set of alleles from each parent. The Examiner also states that the Applicant has not characterized parentals. Applicant respectfully disagrees. In a backcross conversion the deposited material is repeatedly used as the parental line.

The Examiner rejects claims 9-10, 12-13, 15-17, 28-32, 34-40, 42, and 44 and states that the claims "are also broadly drawn to any plant produced by crossing the exemplified inbred line with any of a multitude of non-exemplified plants, or any descendent of the exemplified cultivar obtained by using that cultivar as one parent in a series of undisclosed crosses for an undisclosed number of generations and with undisclosed breeding partners."

The Examiner rejects claims 9, 10, 28, and 29, that claim the F1 hybrid seed and F1 hybrid plant made with PH0R8 as a parent. Applicant notes that a claim to the F1 hybrid made with a deposited inbred was expressly acknowledged without reservation by the United States Supreme Court In *J.E.M. Ag. Supply, Inc. v. Pioneer Hi-Bred Int'l, Inc.*, 60 USPQ 2d 1865,1873 (S.Ct. 2001), when the Supreme Court wrote, "...a utility patent on an inbred plant line protects the line as well as all hybrids produced by crossing that inbred with another plant line." Further, one of ordinary skill in the art would know how to cross PH0R8 with another maize plant. The F1 hybrid seed and plant produced using PH0R8, regardless of the other maize plant used, is identifiable because it will have one set of alleles coming from PH0R8. One of ordinary skill in the art would be able to run a molecular profile on PH0R8 and the F1 hybrid and be able to identify the F1 hybrid as being produced from PH0R8. Seed pericarp tissue, which is solely maternal in origin, can be used to discern the maternal or paternal origin of the allele sets if necessary. See page 16 of Poethig, R.S. 1982. Maize, the plant and its parts. In: W.F. Sheridan (Ed.) Maize for

Biological Research, University of North Dakota Press, Grand Forks, ND. pp. 9-18, submitted as Appendix B.

As stated in the specification on page 15, lines 1-16, there are many laboratory-based techniques available for the analysis comparison and characterization of plant genotype such as Restriction Length Polymorphisms (RFLPs) and Simple Sequence Repeats (SSRs). Such techniques may be used to identify whether or not PH0R8 was used to develop a hybrid. The Applicant also submits to the Examiner the journal article by Berry et al. (2002). This article discusses the probability of identifying the parents of the hybrid by SSR data when neither parent is known and without the use of pericarp analysis. A copy of the article by Berry et al. is attached to this Amendment and Request for Reconsideration as Appendix C. The results of the experiment showed that using 100 SSR loci markers resulted in correct parental ranking of inbreds for 53 out of 54 hybrids. Applicant also points out that any breeder of ordinary skill in the art will know the identity of both parents used to produce a hybrid.

The Examiner rejects claims 15-17. Claims 15 and 16 remain pending and are to methods of developing a maize plant through the utilization of PH0R8. Applicant points out that anyone of skill in the art would know how to utilize these well established breeding methods with PH0R8. Description of such occurs throughout the specification and descriptions can also be found in introductory plant breeding books.

The Examiner rejects claims 40 and 42. Claims 40 and 42 have been amended. New claim 50 has been added. Claim 40 is to the method of producing a first generation PH0R8-derived hybrid maize plant. Applicant believes the patent office has previously indicated this claim scope as allowable and requests that this rejection be withdrawn. Claim 42 is to the method of selfing the first generation hybrid PH0R8 for successive filial generations. This is a basic and well known breeding methodology, and the use of this methodology with PH0R8 is described in the specification on page 20, lines 1 to 15. Claim 50 is to plants derived from claim 42 that have at least 50% of their genetics derived from PH0R8. These claimed plants are clearly described by their method of

production, which requires the use of PH0R8. Such plants must be produced through the use of PH0R8 and the Examiner acknowledges that PH0R8 is clearly identified. Further, Applicant has added the limitation of at least 50% inheritance from the PH0R8 side of its pedigree to further emphasize the significant influence of PH0R8 in the claimed product. Genetic inheritance has been accepted by both courts and governmental agencies as an accurate and reliable means of identification. In paternity cases courts routinely compel genetic testing of putative fathers to establish paternity, and federal law mandates that states have laws requiring that genetic test results be admissible in such cases without the necessity for foundation testimony or other proof. 42 U.S.C. 666(a)(5)(F)(iii)(Supp. V 1999). In such cases, a child will, on average, inherit 50% genetic contribution from each parent. Similarly, the plants produced by the method of claim 42 will also, on average, inherit 50% genetic contribution from each parent.

Applicant requests that the Examiner examine the sufficiency of description of claim 50 with all of its claim limitations, including the limitation that the progeny be produced by the method of claim 42, with the use of PH0R8 and retaining at least 50% genetic contribution from PH0R8. One of ordinary skill in the art would know how to cross PH0R8 to develop an F1 hybrid and also how to self plants derived from the cross with PH0R8. In *Ex parte Parks*, 30 USPQ 2d 1234 (B.P.A.I. 1994), the Board of Appeals stated, "Adequate description under the first paragraph of 35 U.S.C. 112 does not require *literal* support for the claimed invention. Rather, it is sufficient if the originally-filed disclosure would have conveyed to one having ordinary skill in the art that an appellant had possession of the concept of what is claimed." Emphasis added. In *J.E.M. Ag. Supply*, the Supreme Court also acknowledged the value of a newly developed line in further breeding, when it stated that, "...a breeder can use a plant that is protected by PVP certificate to 'develop' a new inbred line while he cannot use a plant patented under §101 for such a purpose." Id. at 1873.

The Examiner cites the Federal Circuit as stating that the written description of an invention "requires a precise definition, such as by structure,

formula, [or] chemical name, of the claimed subject matter sufficient to distinguish it from other materials." *University of California v. Eli Lilly and Co.*, 119 F. 3d 1559, 1568; 43USPQ2d 1398, 1406 (Fed. Cir. 1997). The Applicant has fulfilled this written description requirement through the seed deposit of PH0R8. As described in the specification, lines 1-16 on page 15, the seed deposit allows one of ordinary skill to run a molecular profile of PH0R8. Applicant submits the molecular profile of inbred line PH0R8 in the declaration of Dinakar Bhattramakki attached hereto as Appendix D. Further Applicant amends the specification to include such SSR profile. Such SSR profile is not new matter, as it is an inherent feature of inbred line PH0R8, a representative sample of which has been deposited with the ATCC. For example, see Ex parte Marsili, Rosetti, and Pasqualucci, 214 USPQ 904 (1972), in which the Patent and Trademark Office Board of Appeals held that it was not new matter to amend the structure of a compound when a more refined analytic investigation showed a corrected formula. The Board, relying on well established cases of In re Nathan et al, 51 CCPA 1059, 328 F.2d 1005, 140 USPQ 601 (1964); In re Sulkowski, 487 F.2d 920, 180 USPQ 46 (CCPA 1973); Spero v. Ringold, 54 CCPA 1407, 377 F.2d. 652, 153 USPQ 726 (1967), and Petisi et al. v. Rennhard et al., 53 CCPA 1452, 363 F. 2d 903, 150 USPQ 669 (1966) concluded that the "products described, exemplified and claimed by Appellants inherently had and have now the structure given in the amendment in question. Consequently, the changes made in this amendment do not constitute new matter. Marsili at 906. Similarly, in the present case, inbred line PH0R8 inherently had and still has the SSR marker profile being added. One of ordinary skill in the art can use molecular markers to identify PH0R8, a transgenic version of PH0R8, a backcross conversion of PH0R8 and the F1 plant of the transgenic version and backcross conversion of PH0R8.

The Examiner also stated, in reference to Lilly, that, "the court also concluded that 'naming a type of material generally known to exist, in the absence of knowledge as to what that material consists of, is not a description of that material. Id." This is not the case here. Applicant has created a novel line

and seeks a scope of protection that adequately protects the invention. Applicant believes that the derivatives, variants and closely related progeny easily and routinely created by use of this newly developed line are encompassed within the scope of the invention of the variety itself. These derivatives, variants and closely related progeny derive direct and substantial benefit from Applicant's work and deserve to be included within the scope of the claims. Thus, the issue here is patent scope around what has already been created (and deposited) by Applicant. The fact that the progeny have not been created does not prevent them from being protected in this manner. As stated in MPEP 2163 (3) (a), "An invention may be complete and ready for patenting before it has actually been reduced to practice."

In Enzo vs. Gen-Probe, U.S. State Court of Appeals for the Federal Circuit, 63 USPQ 2d 1609, the court reversed its prior decision regarding the insufficiency of the deposited genetic probes to meet the written description requirement. In so holding, the court stated, " As the deposited sequences are about 850, 8500, and 1300 nucleotides long, ..., there are at least hundreds of subsequences of the deposited sequences, an unknown number of which might also meet the claimed hybridization ratio. Moreover, Enzo's expert, Dr. Wetmur, stated that 'astronomical' numbers of mutated variations of the deposited sequence also fall within the scope of those claims, and that such broad claim scope is necessary to adequately protect Enzo's invention from copyists who could otherwise make minor change to the sequence and thereby avoid infringement while still exploiting the benefits of Enzo's invention. The defendants assert that such breadth is fatal to the adequacy of the written description. On the other hand, because the deposited sequences are described by virtue of a reference to their having been deposited, it may well be that various subsequences, mutations, and mixtures of those sequences are also described to one of skill in the art. We regard that question as an issue of fact...."

The issue of whether the progeny as now claimed satisfies the written description requirement is also an issue of fact. PH0R8 is a unique inbred, as

evidenced by the morphological and physiological traits given in Table 1, pages 17-19, of the application. Routinely used molecular techniques, discussed on page 15, lines 1-16 of the application, can be used to verify whether PH0R8 is within the pedigree of a claimed plant. One of ordinary skill in the art would also know from breeding records if PH0R8 were utilized in the development of a claimed plant.

As stated in the written description guidelines "an applicant shows possession of the claimed invention by describing the claimed invention with all its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention. Possession may be shown in a variety of ways, including...by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention." 1255 Official Gazette 140 (Feb. 5, 2002). Phenotypic traits are used in the text of the specification. Genetic and other molecular profiles may be obtained from the deposit. Once a line is identified as being PH0R8, one of ordinary skill in the art would also easily be able to determine which progeny they develop from that line fall within the scope of the claims.

Within the plant breeding arts, breeders use pedigree as a means to characterize lines in reference to their progenitors. It is unambiguous and easily traceable through breeding records that are maintained by any breeder of ordinary skill in the art. It indicates that a line fewer crosses away from a starting line will be, as a whole, more highly related to the starting line. Thus, the work of the original breeder in developing the starting line will be retained in the closely related progeny. More specifically, traits and linkage groups present in PH0R8 will be retained in progeny that are within one breeding cross from PH0R8. Applicant submits that characterization of the progeny of PH0R8 by virtue of their filial relationship is a clear and acceptable means of identification. Not only are filial descriptions used by breeders to evaluate materials for use in their breeding programs, but it is standard practice within the plant breeding industry for universities and companies that license inbred maize lines to retain a royalty from lines developed through the use of their inbreds. Those royalties are, in

almost all cases, based on the filial relationship between the licensed inbred used in breeding and the progeny line commercialized. This is further evidence that those of ordinary skill in the art of plant breeding describe progeny in terms of pedigree and find it an acceptable means of characterization.

As noted in the specification, the development of an inbred line is a time consuming and labor intensive activity. On average, between 10,000 to 20,000 lines are created and screened in order to develop any maize inbred line for which the Applicant files a patent application. Once developed, the inbred line is useful for two purposes: (1) to make commercial hybrids, and (2) as a source of breeding material for the development of new inbreds that retain the original inbred's desired characteristics. A breeder desiring to make a line with similar traits to PH0R8 would be greatly advantaged by being able to use PH0R8 as starting material. This is because the linked genes arranged through Applicant's breeding efforts, and fixed in PH0R8, can be maintained in the progeny of PH0R8 by a breeder of ordinary skill in the art. The end result is the development of an inbred line with substantial benefit from the Applicant's work.

PH0R8-derived progeny are described by the fact that PH0R8 is utilized in a breeding program to make the PH0R8-derived progeny, PH0R8 gives genetic contribution to the PH0R8-derived progeny, and the genetics of PH0R8 are described by ATCC deposit of PH0R8 seed. By limiting the progeny to one breeding cross away from PH0R8 and by limiting the progeny to those that contain at least 50% of their genetics from PH0R8, the Examiner's concern that the progeny may be only distantly related to PH0R8 is addressed.

Applicant would also like to emphasize that PH0R8 cannot be derived through any other means than through PH0R8 seed and plant, nor can the influence of PH0R8 on the progeny be removed from a line within one outcross of PH0R8. To view this claim as one of breadth ignores an essential limitation of the claim; that only a plant developed through the use of PH0R8 is within the scope of the claim. Such a plant could not be obtained without the use of PH0R8, so the claim would not in any way restrict the work of a breeder that did not in fact use PH0R8. Compliance with the written description requirement is

essentially a fact based inquiry that will "necessarily vary depending on the nature of the invention claimed." Vas-Cath v. Mahurkar, 935 F. 2d 1555 (citing In re DiLeone, 436 F2d. 1404, 1405). Thus, the compliance with the written description requirement must be judged in view of this limited scope of the progeny claims. As amended, the claims are drawn to only a limited scope of progeny, progeny whose existence is the direct result of the use of PH0R8. This is in harmony with the statement in section 2163 of the MPEP that "the written description requirement promotes the progress of the useful arts by ensuring inventions are adequately described in the specification in exchange for the right to exclude." That quid pro quo of patent law has been met by the Applicant in the present case, and to use written description to deny adequate patent protection would be contrary to the stated purpose of the written description requirement.

The Examiner also rejects claims 37-39 under 35 USC § 112, first paragraph. Claims 37-39 have been amended for clarification purposes. Claims 37-39 are directed to growing out an F1 hybrid in which PH0R8 is a parent and searching for PH0R8 inbred seed. Due to the imperfect process of seed production parent seed can sometimes be contained in the hybrid seed bag. This claim covers the method of searching for inbred PH0R8 seed within a bag of hybrid seed. The method is clearly described in the specification on page 5, line 21 through line 7 on page 6. One of ordinary skill in the art can practice such a method without undue experimentation. The Applicant requests that the Examiner withdraw his rejection to claims 37-39.

The Examiner rejects claims 3, 9-13, 15-20, 22, 28-32, 34-40, 42, 44, and 47-49 under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. Applicant traverses the rejection.

With the exception of the arguments pertaining to Hunsperger et. al., Kraft et. al. and Eshed et. al., the Examiner provides nearly identical arguments for the 112, first paragraph, lack of enablement rejection as those provided for the 112, first paragraph, lack of written description rejection addressed above.

To avoid repetition, Applicant respectfully requests that the Examiner consider the arguments made in response to the 112, first paragraph, lack of written description rejection as also applicable to the 112, first paragraph, lack of enablement rejection. In addition, the Applicant directly addresses the arguments raised by the Examiner that relate to Hunsperger, Kraft and Eshed.

The Examiner has cited Hunsperger, Kraft and Eshed and stated that they "teach that it is unpredictable whether the gene or genes responsible for conferring a phenotype in one plant genotypic background may be introgressed into the genetic background of a different plant, to confer a desired phenotype in said different plant." The Examiner states that, "Hunsperger et al teach that the introgession of a gene in one genetic background in any plant of the same species, as performed by sexual hybridization, is unpredictable in producing a single gene conversion plant with a desired trait (see, e.g., column 3, lines 26-46)." Applicant's respectfully disagree that this is what is taught by Hunsperger et al. Hunsperger et al. teaches that a gene that results in dwarfism of a petunia plant can be incorporated into other genetic backgrounds of the petunia species (See column 2, line 67 to column 3, lines 1-4). Hunsperger et al. merely discusses the level of the expression of that gene differed in petunia plants of different genetic backgrounds. Hunsperger et al. succeeded in incorporating the gene into petunia plants of different genetic backgrounds. Therefore, Hunsperger et al. support the fact that one can introgress a specific trait into a recurrent parent through backcross conversion. Applicant's specification provides ample disclosure of starting materials such as maize inbred PH0R8, a discussion of traditional breeding methods, and examples of transgenes and naturally occurring genes that may be used in such methods. Hallauer et al. (1988) on page 472, submitted in the information disclosure statement, state that, "For single gene traits that are relatively easy to classify, the backcross method is

effective and relatively easy to manage." The teaching of Hallauer relates specifically to corn breeding and corn inbred line development, while Hunsperger et al. relates to petunia.

The Examiner goes on to state that, "Kraft et al. teach that linkage disequilibrium effects and linkage drag prevent the making of plants comprising a single gene conversion, and that such effects are unpredictably genotype specific and loci-dependent in nature (see, e.g., page 323)." Applicant disagrees that the article states such points. Kraft et al. make no mention of a plant comprising a single gene conversion. Further, Kraft et al. relates to linkage disequilibrium and fingerprinting in sugar beet, a crop other than maize. Kraft et al. state, on p. 326, first column, "The generality of our results for other crop species needs to be investigated."

It is understood by those of skill in the art that backcross conversions are routinely produced and do not represent a substantial change to a variety. The World Seed Organization, on it's web site, writes, "The concept of an essentially derived variety was introduced into the 1991 Act of the UPOV Convention in order to avoid plagiarism through mutation, multiple back-crossing and to fill the gap between Plant Breeder's Rights and patents." As determined by the UPOV Convention, essentially derived varieties may be obtained for example by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, backcrossing, or transformation by genetic engineering. The commercialization of an essentially derived variety needs the authorization of the owner on the rights vested in the initial variety." International Convention for the Protection of New Varieties of Plants, as amended on March 19, 1991, Chapter V, Article 14, Section 5(c), (emphasis added). A copy of the relevant portion of the UPOV Convention and the World Seed Organization web site is attached as Appendix E.

An example of how one of ordinary skill in the art can transfer a gene conferring a qualitative trait into a variety through backcrossing is demonstrated by the fact that the commercial market now distributes a multitude of products produced in this manner. Such conversion lines are easily developed without

undue experimentation. Poehlman et al. (1995) on page 334, submitted in the information disclosure statement, states that, "A backcross-derived inbred line fits into the same hybrid combination as the recurrent parent inbred line and contributes the effect of the additional gene added through the backcross."

The Examiner goes on to state that, "Eshed et al. teach that in plants, epistatic genetic interactions from the various genetic components comprising contributions from different genomes may effect quantitative traits in genetically complex and less than additive fashion (see, e.g., page 1815). The Applicant would like to point out on page 1816, column 1, lines 1-5 of the Eshed et al. article it states, "Recent studies that detected epistasis of selected QTL in Drosophila (Long et al. 1995), soybean (Lark et al.1995) and maize (Doebley et al.1995; Cockerham and Zeng 1996) did not show a less-than-additive trend." Emphasis added. Applicant also adds that transferring a qualitative trait does not require undue experimentation. Please note Hallauer et al. (1988) on page 472, submitted in the information disclosure statement, which states, "For single gene traits that are relatively easy to classify, the backcross method is effective and relatively easy to manage." In new claim 52, the genes transferred into PH0R8 are now limited to the traits of herbicide resistance, insect resistance, disease resistance, male sterility, and waxy starch.

The Examiner alleges that Applicant's traversal of the art rejection on pages 7-8 of the amendment of 12 August 2002 "admits that outcrossing the exemplified inbred to another undisclosed plant is unpredictable." Applicant did not in any way state that it would be unpredictable to introgress a gene into PH0R8 through backcross breeding techniques or transformation.

In light of the amendments to the claims and the foregoing arguments the Applicant requests reconsideration of the rejection under the first paragraph of 35 U.S.C. 112.

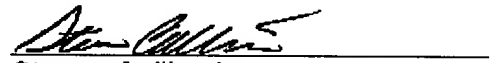
Claims 1, 2, 4-10, 15-16, 21, 23-29, 37-40, 42, and 50-58 are now pending in the application. The amendments made herein do not in any way

change the claim scope which the Applicant believes is allowable but is meant to hasten the issuance of the patent.

CONCLUSION

Applicant submits that in light of the foregoing amendments and the remarks, the claims 1, 2, 4-10, 15-16, 21, 23-29, 37-40, 42, and 50-58 are in condition for allowance. Reconsideration and early notice of allowability is respectfully requested. If it is felt that it would aid in prosecution, the Examiner is invited to contact the undersigned at the number indicated to discuss any outstanding issues.

Respectfully submitted,
Delmar Brenner



Steven Callistein
Reg. No. 43,525
Attorney for Applicant

Steven Callistein
Pioneer Hi-Bred International
7100 NW 62nd Avenue
P.O. Box 1000
Johnston, IA 50131-1000
(515) 254-2823
(515) 224-6883 Facsimile

#10 att

Marker-assisted Selection in Backcross Breeding

S.J. Openshaw

Pioneer Hi-Bred Intl. Inc., P.O. Box 1004, Johnston, IA 50131

S.G. Jarboe[1]

CIMMYT, Lisboa 27, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

W.D. Beavis

Pioneer Hi-Bred Intl. Inc., P.O. Box 1004, Johnston, IA 50131

Abstract. **The backcross breeding procedure has been used widely to transfer simply inherited traits into elite genotypes. Genetic markers can increase the effectiveness of backcrossing by 1) increasing the probability of obtaining a suitable conversion, and 2) decreasing the time required to achieve an acceptable recovery. Simulation and field results indicated that, for a genome consisting of ten 200-cM chromosomes, basing selection on 40 or 80 markers in 50 BC individuals that carry the allele being transferred can reduce the number of backcross generations needed from about seven to three.**

The backcross breeding procedure has been used widely to transfer simply inherited traits into elite genotypes. Usually, the trait being transferred is controlled by a single gene, but highly heritable traits that are more complexly inherited have also been transferred successfully by backcrossing; for example, maturity in maize (Rinke and Sentz, 1961; Shaver, 1976). Today, backcrossing is being used to transfer genes introduced by such techniques as transformation or mutation into appropriate germplasm.

Several plant breeding textbooks give good descriptions of the backcross procedure (Allard, 1960; Fehr, 1987). A donor parent (DP) carrying a trait of interest is crossed to the recurrent parent (RP), an elite line that is lacking the trait. The F_1 is crossed back to the RP to produce the BC_1 generation. In the BC_1 and subsequent backcross generations, selected individuals carrying the gene being transferred are backcrossed to the RP. The expected proportion of DP genome is reduced by half with each generation of backcrossing. Ignoring effects of linkage to the selected DP allele being transferred, the percentage recurrent parent (%RP) genome expected in each backcross generation is calculated as:

$$\%RP = 100\,[1 - (0.5)^{n+1}]$$

where n is the number of backcrosses.

Backcrossing of selected plants to the RP can be repeated each cycle until a line is obtained that is essentially a version of the RP that includes the introgressed allele. After six backcrosses, the expected recovery is >99% (Table 1).

Until recently, discussions of the recovery of the RP genome during backcrossing have emphasized the expected values for %RP shown in Table 1, and have largely ignored the genetic variation for %RP that exists around the expected mean. With the development of genetic markers capable of providing good genome coverage, there has been interest in taking advantage of that variation to increase the efficiency of backcrossing.

Selection for RP marker alleles can increase greatly the effectiveness of backcross programs by allowing the breeder to 1) select backcross plants that have a higher proportion of RP genome, and 2) select backcross individuals that are better conversions near a mapped donor allele being transferred (i.e., select for less linkage drag). Expressed in practical terms, using genetic markers to assist backcrossing can 1) increase the probability of obtaining a suitable conversion, and 2) decrease the time required to achieve an acceptable recovery.

Issues to consider when planning a marker-assisted backcross program include 1) the time advantage of using markers to assist backcrossing, 2) the number of markers needed, and 3) the number of genotypes to evaluate. In this report, we use results from previous literature, computer simulation, and empirical studies to provide some guidelines.

Table 1. Expected recovery of recurrent parent (RP) genome during backcrossing, assuming no linkage to the gene being transferred.

Generation	%RP
F_1	50.0000
BC_1	75.0000
BC_2	87.5000
BC_3	93.7500
BC_4	96.8750
BC_5	98.4375
BC_6	99.2188
BC_7	99.6094

[1]Formerly with Purdue University, West Lafayette, Ind.

Appendix A

Materials and methods

The maize genome was the model for the simulation. The simulated genome contained ten 200-cM chromosomes. Simulation of crossing over was based on a Poisson distribution with a mean of 2.0 ($\lambda = 2$) (Hanson, 1959), which, on average, generated one cross over for every 100-cM length. The simulations reported here assume no interference. Codominant genetic markers were evenly distributed in the genome and sites of the donor gene were randomly assigned to genome locations.

Simulations were conducted with the following parameters:

Number of progeny: 100 or 500.
Backcross generations: BC_1, BC_2, and BC_3.
Number of markers: 20, 40, 80, or 100.
Number selected to form the next BC generation: 1 or 5.

Selection was based on 1) presence of the donor allele and 2) high %RP). %RP was calculated as the average of the (one or five) selected individuals. Values presented are the mean of 50 simulations.

Results

In the computer simulation study, all methods modeled greatly increased the speed of recovering the RP genome compared to the expected recovery with no marker-assisted selection (compare Tables 1 and 2). At least 80 markers were required to recover 99% of the RP genome in just three BC generations (Table 2). Use of at least 80 markers and 500 progeny allowed recovery of 98% RP in just two BC generations. Response to selection was diminished only slightly by spreading the effort over five selections. Using markers, the number of backcross generations needed to convert an inbred is reduced from about seven to three.

By the BC_2 generation, there appears to be no practical advantage to using 500 vs. 100 individuals. If the presence of the donor trait in the backcross individuals can be ascertained before markers are genotyped, then only half the number of individuals indicated in the tables will need to be analyzed.

When a small number of markers are used, they quickly became non-informative; i.e., selection causes the marker loci to became fixed for the RP type before the rest of the genome is fully converted (Table 3; Hospital et al., 1992). This situation was most prominent in the larger populations, where a higher selection intensity placed more selection pressure upon the marker loci. Accordingly, it is of interest to consider how closely the estimation of %RP based on markers reflects the actual genome composition. The combination of estimation of %RP based on fewer markers and subsequent selection tends to bias the estimates upward (compare Tables 2 and 3).

The results from the simulation compare well with real field data. In a typical example, 50 BC_1 plants carrying the gene being transferred were genotyped at 83 polymorphic RFLP loci (note that this corresponds to a population size of 100 unselected plants in Tables 2 and 3). The five best BC_1 recoveries had estimated %RP values of 85.9%, 82.7%, 82.0%, 81.4%, and 81.2%. After evaluating 10 BC_2 plants from each selected BC_1, the best BC_2 recovery had an estimated %RP of 94.6%.

Discussion

The simulations (Table 2; Hospital et al., 1992) and our experience indicate that four markers per 200-cM chromosome is adequate to greatly increase the effectiveness of selection in the BC_1. However, using only four markers per 200 cM will likely make it very difficult to map the location of the gene of interest. Adequate summarization of the data is an important

Table 2. *Percent recurrent parent genome during marker-assisted backcrossing.*

	100 Progeny				500 Progeny			
	No. markers				No. markers			
Generation	20	40	80	100	20	40	80	100
				One selected				
BC_1	84.5	84.5	84.2	88.0	89.9	90.7	90.2	90.5
BC_2	95.0	95.2	95.8	97.2	96.5	97.7	98.5	98.6
BC_3	97.4	97.6	98.9	99.2	97.7	98.3	99.4	99.5
				Five selected				
BC_1	82.9	85.1	84.9	84.7	87.7	88.1	88.9	88.9
BC_2	93.7	95.0	95.8	95.7	95.5	96.8	97.8	97.9
BC_3	97.1	98.3	98.8	98.9	97.3	98.5	99.3	99.3

Table 3. *Estimates of percent recurrent parent genome, based on marker loci.*

	100 Progeny				500 Progeny			
	No. markers				No. markers			
Generation	20	40	80	100	20	40	80	100
				One selected				
BC_2	98.7	97.8	95.6	97.2	100.0	99.1	98.6	98.0
BC_3	100.0	99.8	99.3	99.5	100.0	100.0	99.9	98.2
				Five selected				
BC_2	96.4	96.5	96.2	95.8	100.0	98.5	98.3	98.2
BC_3	99.9	99.8	99.3	99.1	100.0	100.0	99.9	99.8

part of a marker-assisted backcross program. Ideally, the markers used can supply data that can be represented as alleles of loci with known map position. Estimation of %RP, mapping the position of the locus of interest, and graphical display of the results (Young and Tanksley, 1989) are all useful in understanding and controlling the specific backcross experiment being conducted.

It appears that, with the use of genetic markers, the portion of the RP genome that is not linked to the allele being transferred can be recovered quickly and with confidence. The recovery of RP will be slower on the chromosome carrying the gene of interest. A considerable amount of linkage drag is expected to accompany selection for the DP allele in a backcross program. For a locus located in the middle of a 200-cM chromosome, the length of the DP chromosome segment accompanying selection is expected to be 126, 63, and 28 cM in the BC_1, BC_2, and BC_3 generations, respectively (Hanson, 1959; Naveira and Barbadilla, 1992). Our observations support the recommendation of Hospital et al. (1992) that preference be given to the selection for recombinants proximal to the allele of interest, but that selection for recovery of the RP elsewhere in the genome also be considered. This two-stage selection can probably be done quite effectively ad hoc by the breeder once the data is adequately summarized; however, Hospital et al. suggest ways to incorporate the two criteria into a selection index such that each component of selection is assured appropriate weighting.

Use of genetic markers can greatly increase the effectiveness of backcrossing, and they should be used in any serious backcrossing program if resources are available to the breeder.

Literature Cited

Allard, R.W. 1960. Principles of plant breeding. Wiley, New York.

Fehr, W.F. 1987. Principles of cultivar development v.1. Theory and technique. Macmillan, New York.

Hanson, W.D. 1959. Early generation analysis of length of heterozygous chromosome segments around a locus held heterozygous with backcrossing or selfing. Genetics 44:843–847.

Hospital, F., C. Chevalet, and P. Mulsant. 1992. Using markers in gene introgression breeding programs. Genetics 132:1199–1210.

Rinke, E.H and J.C. Sentz. 1961. Moving corn-belt germplasm northward. Ann. Hybrid Corn Industry Conf. 16:53–56.

Shaver, D.L. 1976. Conversions for earliness in maize inbreds. Maize Genet. Coop. Newsltr. 50:20–23.

Young, N.D. and S.D. Tanksley. 1989. Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theo. Applied Genet. 77: 95–101.

2 MAIZE - THE PLANT AND ITS PARTS

R. Scott Poethig
Department of Agronomy, Curtis Hall
University of Missouri
Columbia, MO 65211

One of the greatest deterrents to an appreciation of plant morphology is the terminology used to describe various plant parts. This problem is compounded in the case of maize because of its relatively unusual structure. We all learn that plants have a vegetative body composed of stems, leaves and roots, and that flowers contain sepals, petals, pistils and stamens. Maize, however, has at least three kinds of leaves, two kinds of stems, two kinds of roots, and two kinds of flowers in which glumes, lemmas and paleas take the place of sepals and petals. Fortunately, these parts are arranged in a relatively simple fashion, so the task of mastering maize morphology is not as difficult as it might seem. In this article we will identify some of the most important parts of the maize plant and describe their organization. More detailed descriptions of the developmental morphology of maize have been provided by a number of investigators. Kiesselbach (1949, reprinted 1980) gives a good general picture of maize structure and development. The external morphology and the histology of the vegetative and reproductive shoots have been studied by Bonnett (1948, 1953), Sharman (1942) and Abbe and co-workers (Abbe and Phinney, 1951; Abbe et al., 1951), while the most comprehensive descriptions of the embryogeny are those of Randolph (1936) and Abbe and Stein (1954). A summary of the histology of the corn plant, written by Sass in 1955, has been reprinted in the recent edition of Corn and Corn Improvement (1976).

The organization of the plant body: Maize is a member of the grass family, the Gramineae, and as in all grasses, most of the plant body is leaf tissue (Fig. 1a). To appreciate the general organization of the maize plant it is helpful, therefore, to see it in a leaf-less state (Fig. 1b). Stripped naked, the maize plant is not very impressive. Its main stem, or culm, is a slender, segmented shaft similar to a stalk of bamboo or sugarcane. The enlarged joints along the stem, the nodes, mark the points of leaf attachment; the stem segment between nodes is called the internode. Each node bears a single leaf in a position opposite that of the neighboring leaf, giving the plant two vertical rows of leaves in a single plane (Fig. 1a; 2). This so-called distichous phyllotaxy is typical of all leaf-like appendages, wherever they occur on the plant.

Maize has unisexual, rather than bisexual flowers. Male (staminate) flowers are located at the apical tip of the main stem in the tassel, a branched inflorescence. Female (pistillate) flowers are found in one to several compact ears, located on the ends of short branches near the middle of the stem (Fig. 1b; 2).

This partitioning of male and female flowers in separate structures distinguishes maize from other cereals and is one of the principal reasons that its genetics has been so conveniently explored. Making controlled pollinations in maize requires little more effort than that involved in placing a bag over the tassel and ear shoot. To perform a controlled pollination in rice, wheat, barley and other cereals, it is necessary to emasculate each

Appendix B

flower used as a female parent, an especially tedious job when each flower yields only one seed.

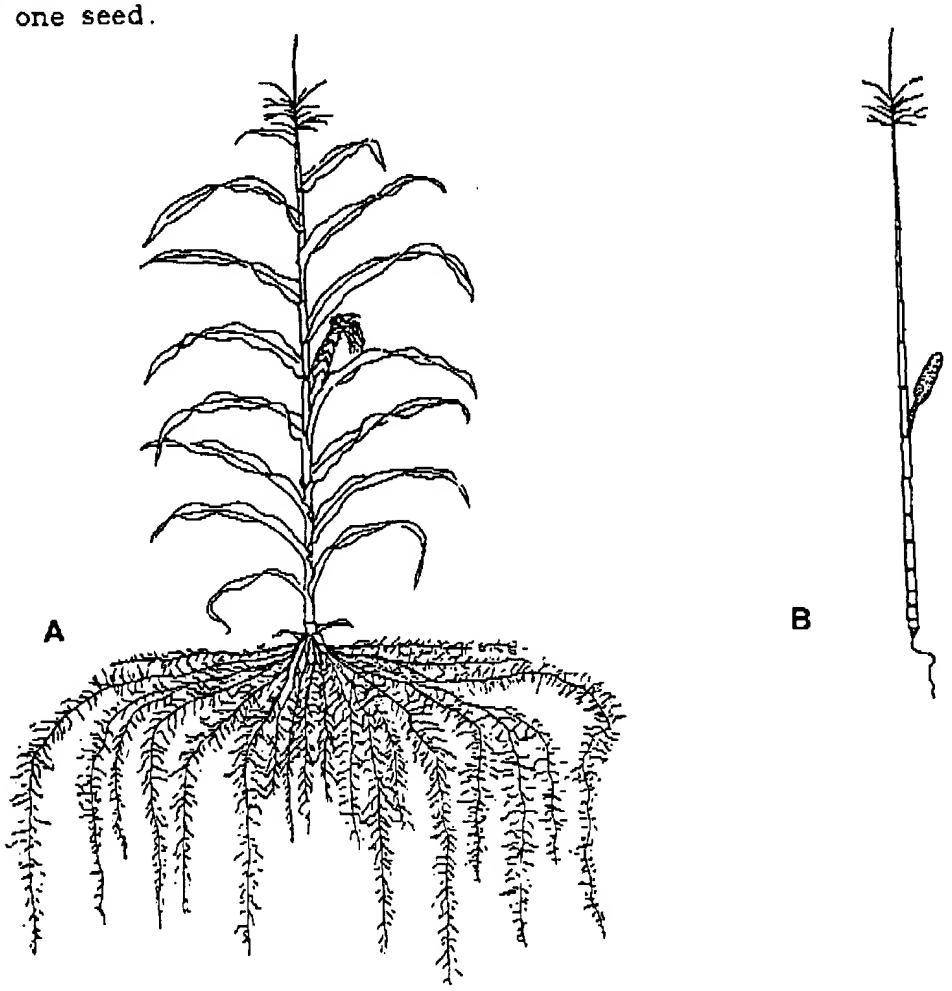


Figure 1. a) Mature maize plant (after Kiesselbach, 1949). b) Mature maize plant drawn without leaves and adventitious roots. The apical end of the main stem (culm) terminates in the tassel, while the basal end terminates in the primary root (radicle). The ear shoot arises from an internode near the center of the culm.

Maize also differs from closely related species in that it has relatively few branches. Only the lower 10 to 12 internodes of the stem produce branch primordia, and most of these remain suppressed. Above-ground primordia develop into ear shoots, while those located at subterranean internodes develop into tillers--branches identical in structure to the main stem. Commercial hybrids (except sweet corns) generally tiller very little, and typically produce a single viable ear shoot. In contrast, some "varieties" may have several large tillers and may produce 2 ears on the main stem and some ears on tillers.

The stem: During the first four weeks after germination, the growing point of the stem lays down all the nodes and internodes of the plant and then differentiates into a tassel. At the time of tassel formation the stem is not more than 3-4 inches tall, even though the plant may be 3-4 feet in

height (Fig. 3). Subsequently, the stem begins to elongate rapidly, with most of the growth occurring at the base of the internodes. The lowermost 6-8 internodes do not participate in this growth, however, and remain below ground where they produce the root system and tillers. These subterranean internodes taper sharply towards the base of the stem, forming a distinctive region, the crown (Fig. 1b). The stem is thickest a few inches above ground, and tapers gradually towards the tassel. All the internodes from the top ear downward have a distinct groove associated with the axillary bud at the base of the internodes; internodes above the ear lack axillary buds and are smoothly cylindrical.

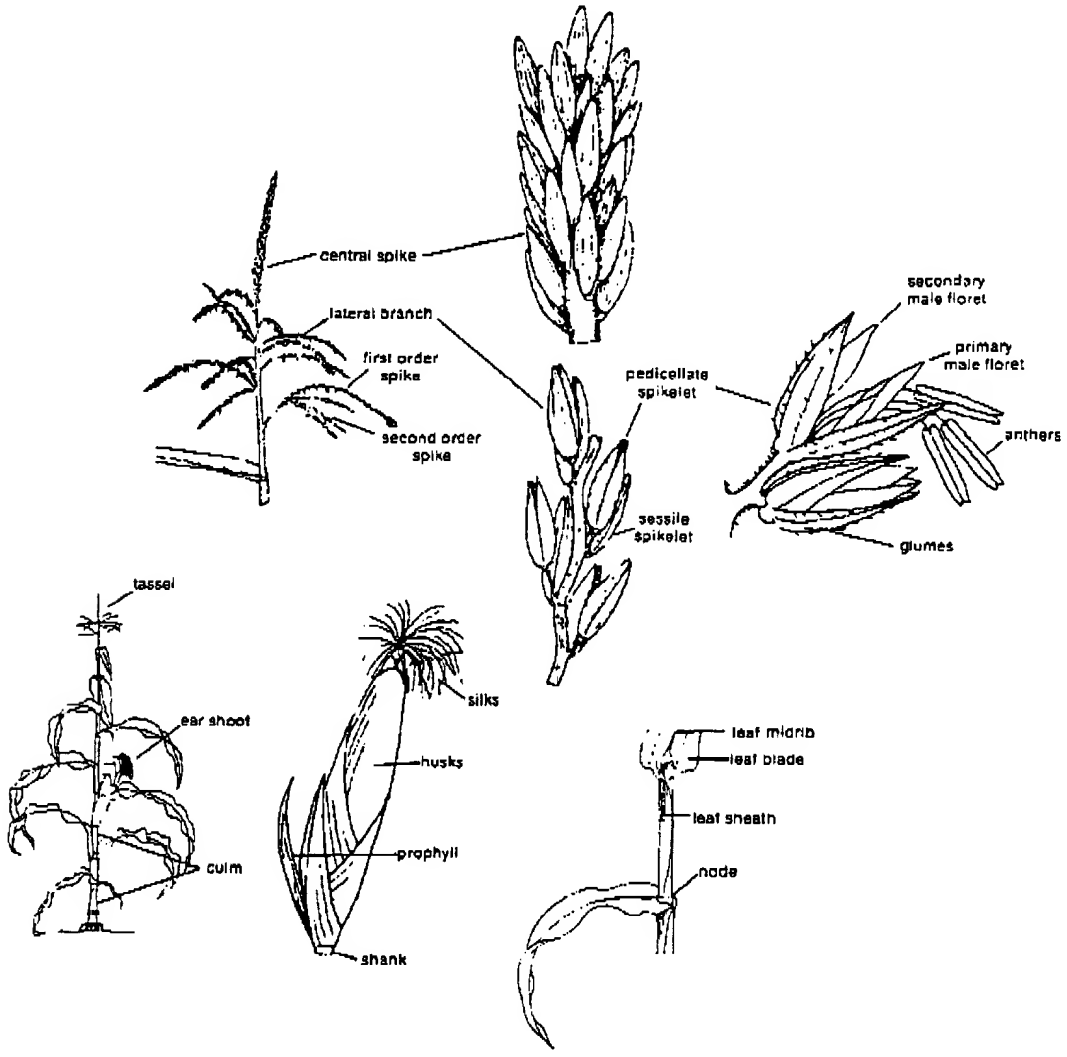


Figure 2. The major parts of the maize plant. Drawings in part from P. Weatherwax in Corn and Corn Improvement, 1955, and E. D. Styles et al. in Can. J. Genet. Cytol. 15:59, 1973; figure assembled by M. M. Johri and E. H. Coe.

The stem of an ear shoot, called the shank (Fig. 2), differs from the main stem in being relatively short in most strains. In addition, the internodes of the shank are variable in number, irregular in shape and size, and tend to have a crinkled rather than smooth surface. Secondary ear shoots commonly occur on the shank of several types of maize, but are rare in most commercial strains unless fertilization of the apical ear is prevented.

Figure 3. A four week old plant (approximately 3 feet tall) in which the stem apex has differentiated into a tassel. As shown on the right, the stem is still relatively short at this stage.

The tassel: The tassel, located at the top of the culm, consists of a series of large branches (spikes) covered with numerous, small flower-bearing branches (spikelets: Fig. 2) Each branch point on a spike bears two spikelets, one on a long stem (pedicellate), the other on a short stem (sessile) (Fig. 4a). Each of these spikelets, in turn, produces two functional florets. Although tassel florets contain both stamens and a pistil, the pistil normally degenerates soon after it is initiated, making the floret functionally male. However, pistils will develop at the base of the tassel under some environmental and physiological conditions, and are quite common on tillers.

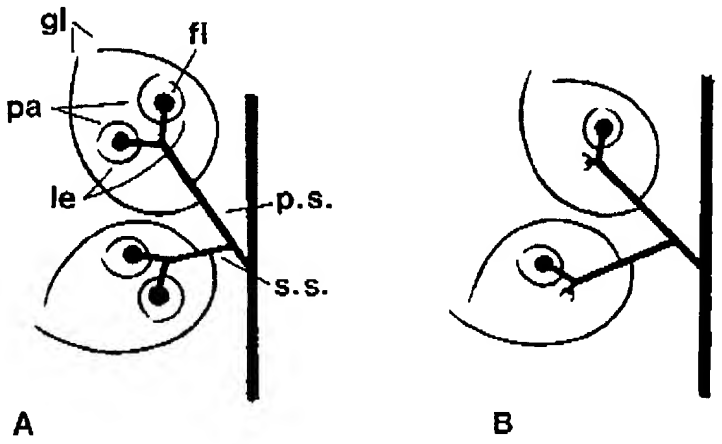


Figure 4. Schematic drawing of a pair of tassel spikelets (A) and a pair of ear spikelets (B). Note that the lower floret in the ear spikelet aborts early in development. p.s. - pedicellate spikelet; s.s. - sessile spikelet; gl - glumes; le - lemma; pa - palea; fl - floret.

Surrounding both florets on a spikelet are 2 leaf-like scales called glumes (Fig. 2; 4a). Within the glumes, each floret is individually enclosed in another pair of scales, one located adjacent to the glume (the lemma), the other located between the two florets (the palea) (Fig. 4a). At anthesis, these scales are forced apart by the swelling of conical structures (lodicules) at the base of the 3 stamens, and the filamentous base of the stamens elongates, forcing the anthers out of the flower (Fig. 2). As they dangle downwards, the anthers shed pollen from openings at their tip.

Pollen grains are the multicellular products of the haploid microspores that result from the meiosis of a microspore mother cell (microsporocyte). Meiosis takes place in the anther before the tassel emerges from the leaf sheaths. After meiosis, the 4 resulting haploid microspores separate from each other, and each forms a thick wall. Shortly before shedding, each microspore undergoes two mitotic divisions. The first division is asymmetric, and produces a relatively large vegetative cell and a smaller generative cell. In the second division, the generative cell divides to form two sperm cells.

The ear: The ear is morphologically similar to the tassel, although this resemblance is obscured by differences in the relative size of their parts. The crucial difference between them is, of course, that the tassel contains male flowers, and the ear bears female ones. This difference is due simply to the fact that during the formation of an ear floret, stamen primordia are arrested at an early stage in their development, while the pistil develops fully. Each functional ear floret has a single ovary, which terminates in an elongated style, or silk (Fig. 5). Within the ovary is a single embryo sac. The embryo sac is the product of one of the four haploid cells resulting from the meiosis of the megaspore mother cell. While its three sister cells degenerate, the nucleus of this cell divides three times to produce 8 haploid nuclei within a common cytoplasm (the embryo sac). Two of these nuclei (polar nuclei) migrate to the center of the embryo sac where they become closely associated. The three nuclei remaining at the base of the embryo sac

subsequently undergo cellularization to form the egg cell and two synergids, while the 3 nuclei at the tip of the embryo sac proliferate to form 24-48 antipodal cells.

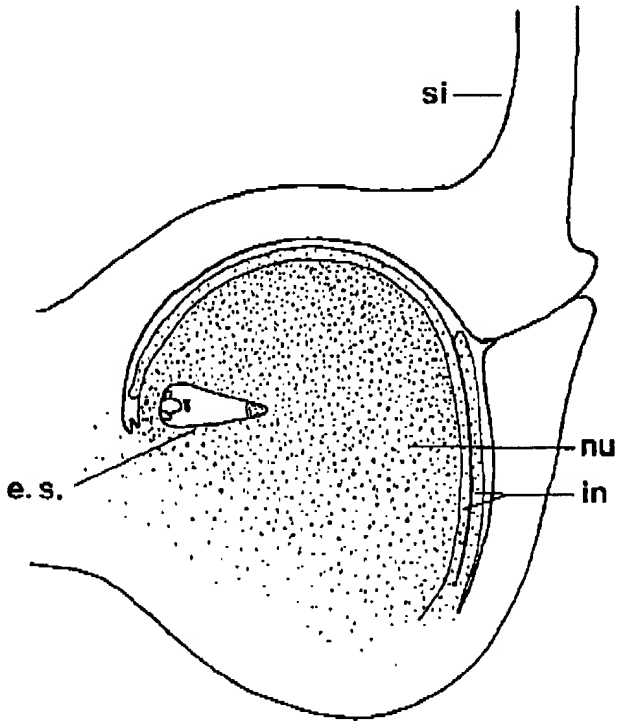


Figure 5. Radial longitudinal section of an ovary with an unfertilized embryo sac (after Randolph, 1936). Upon fertilization, the nucellus is digested by the expanding embryo sac and the tissue surrounding the nucellus is transformed into the pericarp. si - silk; e.s. - embryo sac; nu - nucellus; in - integuments.

The ear also differs from the tassel in that it has no major lateral branches. Its thick, lignified axis, the cob, is homologous to the central spike of the tassel. As in the tassel, ear spikelets come in pairs, but in the ear they are equal in size and only one of the florets in each spikelet is functional (Fig. 4b). An ear therefore has an even number of parallel rows of equally sized kernels equal to the number of spikelets on the cob. The number of rows (or ranks) of kernels ranges from 4 to 30.

The glumes, lemmas and paleas of the ear spikelets are readily visible in an unfertilized ear, but are soon obscured by the enlargement of the ovary after fertilization. In a mature ear these structures are represented by the chaff that adheres to the cob and the base of the kernel after it is shelled.

The leaf: Maize produces three kinds of vegetative leaves: foliar leaves, husk leaves and prophylls. A foliar leaf is located at each of the nodes on the main stem, husk leaves are located on the shank of the ear shoot, and prophylls are found at the base of the shank between the ear shoot and the stem (Fig. 2).

The foliar leaf has two distinct parts--the blade, a flat portion extending away from the stem, and the sheath, a basal part that wraps tightly around the stem (Fig. 2). Internally, the blade consists of a spongy network of cells traversed by a series of parallel, longitudinal veins. This flexible lamina is supported by the midrib, a thickened, translucent structure located in the center of the leaf. The sheath is thicker and more rigid than the blade, possesses fewer longitudinal veins, and lacks a prominent midrib. The sheath completely encircles the internode above the node to which it is attached and may extend the entire length of that internode.. During the early development of the plant, the leaf sheaths provide most of the mechanical support necessary to keep the stem upright. At the boundary between the blade and the sheath there is a distinct hinge of translucent tissue. In this region, the leaf blade and leaf sheath narrow sharply, forming an indentation in the leaf margin. The wedge of translucent tissue adjacent to this indentation is known as the auricle. The ligule is the thin collar of filmy tissue located on the inside of the hinge.

The husk leaves surrounding the ear are usually considered modified leaf sheaths, with vestiges of the blade portions occasionally present. In some strains husk leaves develop a prominent ligule and leaf blade. In contrast to the leaf sheath, husk leaves are relatively thin and flat. Each husk leaf is attached to a unique node on the shank, and all but a few upper ones are arranged distichously.

Located between an ear shoot and the stem, the prophyll looks superficially like a husk leaf, but is distinguished by having two keels (midribs) and a split apex. These features suggest that the prophyll arose evolutionarily from the fusion of two foliar leaves. The homology of the prophyll is still controversial, however. Galinat (1959), for example, considers the prophyll one of the basic units of maize morphology, the others being the internode, leaf and axillary bud.

The root: More is known about the growth, cell biology, physiology and anatomy of the primary maize root, or radicle, than perhaps any other organ of the plant. Its histological structure, described by Sass (1976) and Kiesselbach (1949), is typical of roots in general. The apex of the root is sheathed in a loose network of root cap cells. Immediately behind the apex is a zone of cell division and elongation, beyond which root hairs are initiated. Larger lateral roots arise at varying points behind the zone of root hair formation. Cell division is restricted to the apical 3 mm of the root, and occurs at a maximal rate 1.25 mm behind the apex. The zone of elongation extends 8 mm behind the apex, the rate of elongation being maximal 4 mm from the tip (Erickson and Sax, 1956). Those interested in using the root for physiological or cell cycle studies should consult Silk and Erickson (1979; 1980) and Green (1976) for an analysis of the growth parameters that must be taken into consideration in such studies.

The primary root represents the basal end of the plant axis, which in maize and other grasses contributes relatively little to the ultimate root system (compare Fig. 1a and b). Most of the root system consists of adventitious roots produced by the basal-most internodes of the stem. The primordia of a few adventitious roots are normally present in the embryo, and these emerge soon after germination. New root primordia are subsequently initiated at the base of all subterranean internodes, and also appear

at 2 or 3 above-ground internodes after the stem has elongated. Subterranean adventitious roots are sometimes called crown roots, while those initiated above ground are known as brace roots.

Adventitious roots grow horizontally for several feet before turning downwards. As a result, the root system of a single plant often covers a region 6-8 feet in diameter, while the depth of the root system may be as much as 6 feet. As it grows, the root branches profusely in the region behind the apex, forming both secondary roots and unicellular root hairs. The total length of root system of a mature plant has been estimated to be 6 miles.

The kernel: The events surrounding the process of fertilization have been described by Miller (1919), Kiesselbach (1949) and Pfahler (1975); unfortunately, ultrastructural information about this phenomenon is still unavailable.

The silk is receptive to pollen along its entire length. Within 5 minutes after a pollen grain lands on a silk it sends out a tube which penetrates the silk and grows downward towards the ovary. During this process the vegetative nucleus and the two sperm cells migrate to the tip of the pollen tube where they remain throughout its growth. Upon reaching the embryo sac, 12 to 24 hours after germination, the end of the pollen tube bursts, releasing the two sperm. One sperm nucleus fuses with the two polar nuclei in the center of the embryo sac to form a triploid cell that gives rise to the endosperm. The other sperm nucleus fuses with the egg nucleus to form the zygote. As often as 2% of the time the polar nuclei and the egg nucleus are fertilized by sperm from different pollen grains, with the extra sperm nuclei being somehow lost (Sarkar and Coe, 1971). This phenomenon, called heterofertilization, can lead to a non-correspondence between the genotype of the endosperm and embryo when the male parent is heterozygous.

The development of the kernel following fertilization has been described in detail by Randolph (1936). We will only note here that this process takes 40-50 days and is accompanied by a 1400-fold increase in the volume of the embryo sac. The growth of the embryo and the accumulation of food reserves in the endosperm is completed by about day 40, and the remaining 10-20 days is spent maturing and drying.

A mature kernel has three major parts: the pericarp, endosperm and embryo (Fig. 6). The pericarp, the tough transparent outer layer of the kernel, is derived from the ovary wall and is therefore genetically identical to the maternal parent. The endosperm and embryo represent the next generation.

The endosperm makes up about 85% of the weight of the kernel and is the food source for the embryo for several days after it germinates. This food takes the form of intracellular starch grains and protein bodies, and is concentrated to varying degrees in different parts of the endosperm (Duvick, 1961). In flint-type kernels the concentration of starch and protein bodies is higher around the periphery of the endosperm than in the center, giving the endosperm a hard, corneous external layer, and a soft, granular center. In dent kernels, the granular tissue extends to the crown of the endosperm so that it collapses upon drying and produces a distinct indentation. These two traits are polygenic in their inheritance and are

characteristic of specific races of maize. Other common endosperm traits, such as sugary, floury or shrunken, are single gene mutations and can exist in either a flint or dent background.

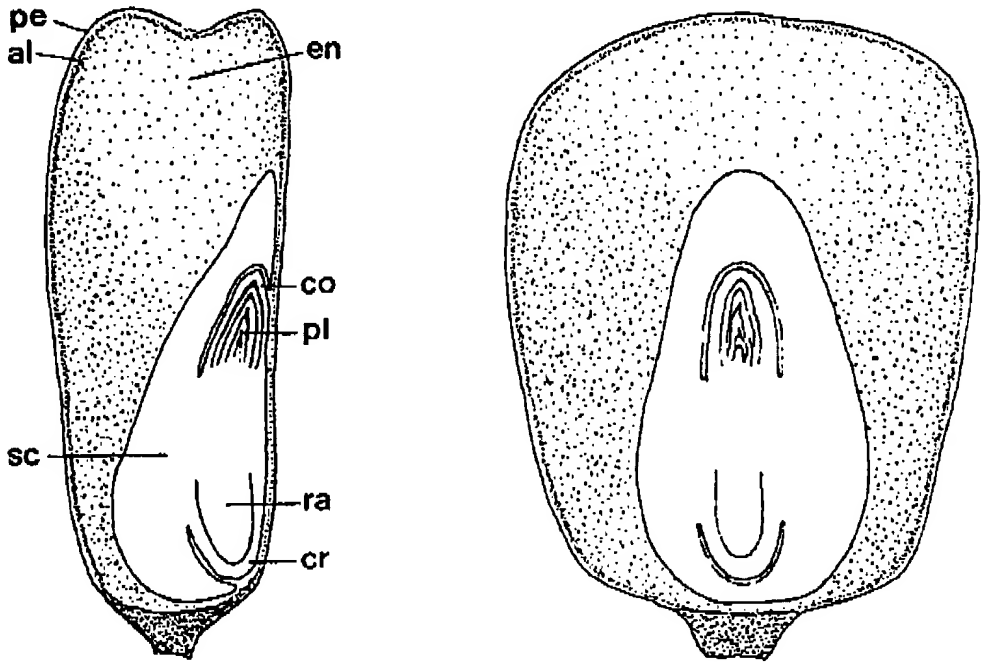


Figure 6. Longitudinal sectors of a mature dent kernel, taken perpendicular (left) and parallel (right) to the upper face of the kernel (after Kiesselbach, 1949). pe - pericarp; en - endosperm; al - aleurone; sc - scutellum; co - coleoptile; pl - plumule; ra - radicle; cr - coleorhiza.

Much of our understanding of gene action in maize is based on the analysis of genes affecting the pigmentation of the external layer of the endosperm, the aleurone. This specialized single cell layer is the only part of the endosperm capable of becoming intensely pigmented. Internal endosperm cells may be either yellow or white.

The embryo is located on the broad side of the kernel facing the upper end of the ear, beneath a thin layer of endosperm cells. Most of the tissue in the embryo is part of the scutellum, a spade-like structure concerned with digesting and transmitting to the germinating seedling the nutrients stored in the endosperm. The shoot and root axis are recessed in the outer face of the scutellum. In a mature kernel, the shoot (plumule) has 5 to 6 leaf primordia that are arrested at successive stages of development (Abbe and Stein, 1954). Surrounding the shoot is a cylindrical structure called the coleoptile. Upon germination, the coleoptile elongates until it is above ground and is then ruptured by the more rapid expansion of the rolled leaves within it. The root is enclosed in a sheath of tissue called the coleorhiza. Unlike the coleoptile, the coleorhiza does not elongate very much, and gives way to the radicle as soon as it emerges from the seed.

References

Abbe, E. C. and B. O. Phinney. 1951. The growth of the shoot apex in maize: external features. Amer. J. Bot. 38:737-744.

Abbe, E. C., B. O. Phinney and D. F. Baer. 1951. The growth of the shoot apex in maize: internal features. Amer. J. Bot. 38:744-751.

Abbe, E. C. and O. L. Stein. 1954. The growth of the shoot apex in maize: embryogeny. Amer. J. Bot. 41:285-293.

Bonnett, O. T. 1948. Ear and tassel development in maize. Missouri Bot. Gard. Ann. 35:269-287.

Bonnett, O. T. 1953. Developmental morphology of the vegetative and floral shoots of maize. Bull. 568, Agric. Exp. Sta., U. of Illinois.

Duvick, D. N. 1961. Protein granules of maize endosperm cells. Cereal Chem. 38:374-385.

Erickson, R. O. and K. B. Sax. 1956. Elemental growth rate of the primary root of Zea mays. Proc. Amer. Phil. Soc. 100:487-498.

Galinat, W. C. 1959. The phytomer in relation to floral homologies in the American Maydeae. Bot. Mus. Leaflets, Harvard U. 19:1-32.

Green, P. B. 1976. Growth and cell pattern formation on an axis: critique of concepts, terminology and modes of study. Bot. Gaz. 137:187-202.

Kiesselbach, T. A. 1949. The structure and reproduction of corn. Res. Bull. 161, Agric. Exp. Sta., U. of Nebraska College of Agric. Reprinted 1980, U. Nebraska Press.

Miller, E. C. 1919. Development of the pistillate spikelet and fertilization in Zea mays L. J. Agric. Res. 18:255-265, with 14 plates.

Pfahler, P. L. 1978. Biology of the maize male gametophyte. In: Maize Breeding and Genetics D. B. Walden, ed., John Wiley and Sons, Inc.

Randolph, L. F. 1936. Developmental morphology of the caryopsis in maize. J. Agric. Res. 53:881-916.

Sarkar, K. R. and E. H. Coe; Jr. 1971. Analysis of events leading to heterofertilization in maize. J. Hered. 62:118-120.

Sass, J. E. 1976. Morphology. In: Corn and Corn Improvement G. F. Sprague, ed., Amer. Soc. Agronomy, Madison.

Sharman, B. C. 1942. Developmental anatomy of the shoot of Zea mays L. Ann. Bot. N. S. 6:245-282.

Silk, W. K. and R. O. Erickson. 1979. Kinematics of plant growth. J. Theor. Biol. 76:481-501.

Silk, W. K. and R. O. Erickson. 1980. Local biosynthetic rates of cytoplasmic constituents in growing tissue. J. Theor. Biol. 83:701-703.

Assessing Probability of Ancestry Using Simple Sequence Repeat Profiles: Applications to Maize Hybrids and Inbreds

Donald A. Berry,*[,1] Jon D. Seltzer,[†] Chongqing Xie,[‡] Deanne L. Wright[‡] and J. Stephen C. Smith[‡]

*The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, [†]Third Wave Technologies, Inc., Madison, Wisconsin 53719 and [‡]Pioneer Hi-Bred International, Inc., Johnston, Iowa 50131

Manuscript received July 24, 2001
Accepted for publication March 11, 2002

ABSTRACT

Determination of parentage is fundamental to the study of biology and to applications such as the identification of pedigrees. Limitations to studies of parentage have stemmed from the use of an insufficient number of hypervariable loci and mismatches of alleles that can be caused by mutation or by laboratory error and that can generate false exclusions. Furthermore, most studies of parentage have been limited to comparisons of small numbers of specific parent-progeny triplets thereby precluding large-scale surveys of candidates where there may be no prior knowledge of parentage. We present an algorithm that can determine probability of parentage in circumstances where there is no prior knowledge of pedigree and that is robust in the face of missing data or mistyped data. We present data from 54 maize hybrids and 586 maize inbreds that were profiled using 195 SSR loci including simulations of additional levels of missing and mistyped data to demonstrate the utility and flexibility of this algorithm.

DETERMINATION of parentage is fundamental to the study of reproductive and behavioral biology. The increasing availability of highly discriminant genetic markers for many diverse species provides the potential to uniquely characterize individuals at numerous loci and to unambiguously resolve parentage where genealogical relationships are unknown, in error, or in dispute.

Identification of parent-progeny relationships in wild populations of animals and plants provides insights into the success of various reproductive strategies (Ellstrand 1984; Smouse and Meagher 1994; Alderson *et al.* 1999) and has allowed for the implementation of management programs to conserve genetic diversity (Miller 1975; Rannala and Mountain 1997). The association of pedigree with physical appearance or performance in domesticated animals and plants allows parents that have contributed favorable alleles for desirable traits through selective breeding programs to be identified (Bowers and Meredith 1997; Sefc *et al.* 1998; Vankan and Faddy 1999). These applications of associative genetics facilitate further progress in genetic improvement through breeding. Establishment of parentage is also useful to secure legal rights of guardianship in humans, to help protect intellectual property in plant varieties, to validate breed pedigrees of domesticated animals, to protect stocks of fish, and to identify provenance of meat that is available in supermarkets (Gotz and Thaller 1998; Primmer *et al.* 2000; White *et al.* 2000).

Most studies of pedigree have utilized exclusion analysis where the molecular marker genotypes of either one or a restricted number of potential triplets of offspring and putative parents are compared. Often the identity of the mother is not in question; the maternal profile is subtracted from that of the offspring and the deduced paternal profile is then compared with candidate father genotypes (Ellstrand 1984; Hamrick and Schnabel 1985). Individuals who could not have contributed the paternal genotype are excluded; the remainder are possible parents. Nonpaternity in humans is generally declared only on the basis of exclusions exhibited by at least two unlinked and independent loci. This criterion of exclusion reduces the likelihood of a false declaration of nonpaternity on the basis of marker results that are actually due to mutation within the phylogeny. Bein *et al.* (1998) show that evidence of nonpaternity should require exclusions at loci on different chromosomes to avoid erroneous conclusions that would be made due to nondisjunction at meiosis leading to uniparental inheritance. A requirement for at least three independent exclusions to declare nonpaternity in humans has also been instituted (Gunn *et al.* 1997). In studies of natural populations of animals or plants where numerous parent-progeny triplets are examined it is usual to accept a single exclusionary event as evidence of nonpaternity (Marshall *et al.* 1998). Paternity testing has been extended to situations where DNA from either parent is unavailable. For example, paternity can still be established in circumstances where the putative father is deceased but his parents are still alive (Helminen *et al.*

[1]*Corresponding author:* Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Box 447, Houston, TX 77030-4009. E-mail: dberry@mdanderson.org

Genetics 161: 813–824 (June 2002)

Appendix C

1991; Bucket *et al.* 1992). Chakraborty *et al.* (1994) demonstrate that paternity can be determined in cases where the mother is unavailable for testing. Lang *et al.* (1993) partially reconstructed the DNA profile of a missing crocodile parent using profiles of the mother and progeny.

Chakraborty *et al.* (1988) and Smouse and Meagher (1994) report that reliance upon exclusion alone has usually failed to unambiguously resolve paternity. Limitations have stemmed from the use of an insufficient number of independent hypervariable loci. Other statistical methods are therefore required to calculate the likelihood of paternity for each nonexcluded male (Berry and Geisser 1986; Meagher 1986; Meagher and Thompson 1986; Thompson and Meagher 1987; Devlin *et al.* 1988; Berry 1991). Marshall *et al.* (1998) draw attention to the quality of data that is encountered practically in genotypic surveys. Maternal genetic data may or may not be available, data may be absent for some candidate males, data may be missing for some loci in some individuals, null alleles exist, and typing errors occur. Reconstructing or validating the pedigrees of varieties of cultivated plants often provides additional challenges because their phylogenies can reveal apparent exclusions that masquerade as non-Mendelian inheritance. For example, apparent exclusions can occur in circumstances where an individual is used as a parent prior to completion of the inbreeding process. The development of parent and progeny then continue on parallel but separate tracks thereby allowing the possibility that alleles that are subsequently lost through inbreeding in the parent can still become fixed in the progeny. It is also possible to create many offspring from a single mating and to use the same parent repeatedly in "backcrossing." Therefore, many individual inbred lines, varieties, or hybrids can be highly related. In consequence, there are numerous (and often very similar) pedigrees. The effective number of marker loci that can discriminate between alternate pedigrees is proportionally reduced as parents are increasingly related. Consequently, inbred lines can be more similar to one or more sister or other inbreds than those inbreds are to one or both of their parents.

It has not been usual to search among hundreds of individuals to identify the most probable maternal and paternal candidates for a specific progeny. Most studies of parentage are in circumstances where there is *a priori* information for at least one of the parents (usually the maternal parent). Limited availability of marker loci and the lack of very high-throughput genotyping systems offering inexpensive datapoint costs may have focused research on studies that involve relatively few individuals and where there is at least some *a priori* indication of parentage. Studies that have been conducted without *a priori* information on parentage include species where reproductive behavior renders identification of the maternal parent difficult or impossible. Examples include those undertaken on birds that practice brood parasitism (Alderson *et al.* 1999) or extra-pair copulation (Wetton *et al.* 1992) or on species such as the wombat that are difficult to observe in the wild (Taylor *et al.* 1997).

Two circumstances favor a revised approach to the statistical analysis of pedigree. First, molecular marker technologies are rapidly developing and will allow numerous loci to be typed for thousands of individuals rapidly and inexpensively. A greater number and diversity of larger-scale studies of pedigree can be expected within the plant and animal kingdoms including individuals in which there is no prior knowledge of pedigree. A larger number of markers mean a greater chance for errors. Therefore, the second circumstance follows: Procedures that are efficient and robust in the face of apparent exclusions, missing data, and laboratory error are required.

The purpose of this article is to describe and evaluate a methodology that can be used to quantify the probability of parentage of hybrid genotypes. We focus on parentage because it is the primary focus of published literature and it is the easiest level of ancestry to understand. The method is robust in the face of mutation, pseudo-non-Mendelian inheritance (apparent exclusions) due to residual heterozygosity in parental seed sources, missing data, and laboratory error. The methodology has a number of advantages: (i) It can accommodate large datasets of possible ancestors (hundreds of inbreds or hybrids each profiled by >100 marker loci), (ii) it does not require prior knowledge about either parent of the hybrid of interest, (iii) it does not require independence of the markers, and (iv) it can successfully discriminate between many highly related and genetically similar genotypes. We demonstrate the effectiveness of this approach to identify inbred parents of maize (*Zea mays* L.) hybrids using simple sequence repeat (SSR) marker profiles for 54 maize hybrids together with their parental and grandparental genotypes included among a total of 586 inbred lines. The methodology is applicable to the investigation of parentage for all progeny developed from parental mating without subsequent generations of inbreeding.

MATERIALS AND METHODS

Algorithm: Consider an index hybrid whose parentage is unknown or in dispute. Inbreds in an available database are possible ancestors of the hybrid. The objective is to find the probabilities of closest ancestry for each inbred on the basis of information from SSRs from the index hybrid and the inbreds. There is no reason to trim the database by removing inbreds thought to be unrelated to the index hybrid because their lack of relationship will be discovered.

Consider a pair of possible ancestors, inbred i and inbred j. There is nothing special about this particular pair as all pairs will be treated similarly. The process involves calculating the probability that inbreds i and j are in the hybrid's ancestry, repeating this for all pairs of inbreds in the database.

The basis of the algorithm is Bayes' rule (*e.g.*, Berry 1991, 1996). Let $P(i, j|\text{SSRs})$ stand for the (posterior) probability that i and j are ancestors of the index hybrid given the information from the various SSRs. Let $P(i, j)$ stand for the unconditional (or prior) probability of the same event. Finally, $P(\text{SSRs}|i, j)$ is the probability of observing the various SSR results if in fact i and j are ancestors. Bayes' rule says

$$P(i, j|\text{SSRs}) = P(\text{SSRs}|i, j) \times P(i, j) \div \sum\{P(\text{SSRs}|u, v) \times P(u, v)\},$$

where the sum in the denominator is over all pairs of inbreds, indexed by u and v. $P(\text{SSRs}|i, j) \times P(i, j)$ is one of the terms in the denominator. (To compute the denominator in the above expression, fix a particular order to the inbreds in the database and take $u < v$ in expressions involving the pair (u, v). If there are 586 inbreds, for example, then the number of pairs and the number of terms in the denominator is $586(587)/2 = 171{,}991$.) Inbreds i and j may be parents or grandparents or other types of relations or bear no relationship at all to the hybrid. If there are more than two ancestors in the database, such as both parents and all four grandparents, then the possible pairs involving these ancestors will generally have the highest posterior probabilities. If the hybrid's true parents are in the database, then as a pair they will typically have the highest overall posterior probability. If both i and j happen to be related to one particular parent of the hybrid, then as a pair their posterior probability will be low because they will not usually account for many of the alleles that are contributed by the other parent of the hybrid.

We will make the "no-prior-information" assumption that $P(u, v)$ is the same for all pairs (u, v). This implies that this factor is cancelled from both numerator and denominator in the above expression, giving:

$$P(i, j|\text{SSRs}) = P(\text{SSRs}|i, j) \div \sum P(\text{SSRs}|u, v).$$

The problem is then to calculate a typical $P(\text{SSRs}|i, j)$. Assume inbreds i and j are both ancestors. We calculate the probability of observing the resulting hybrid under this assumption. We make no assumptions about relationships among the various inbreds. Other possible ancestors will be considered implicitly in the calculation by allowing their alleles to be introduced through breedings with i and j. However, the nature of such breedings is not specified. Suppose inbred i's alleles are (a, b). Each descendant of inbred i receives one of these two alleles or not. An immediate descendant receives one with probability 1 (barring mutations). A second generation descendant receives one of them with probability 0.5. And so on. Since degree of ancestry (if any) is unknown, we label the actual probability of passing on one of these alleles to be P. Similarly, an allele from inbred j has been passed down to the hybrid or not, and the probability of the former is P. In the following, P will be taken to equal 0.50, although we will also consider $P = 0.99$ in some of the calculations.

Assuming $P = 0.50$ is consistent with the closest ancestors in the database being grandparents. However, we are not interested in grandparents *per se*. If the closest ancestors in the database were parents, then as indicated above P should equal 1 (ignoring mutations and laboratory errors). Our primary concern is when the parents are not in the database. In this case P is no greater than 0.50. Assuming $P = 0.50$ is robust over the middle range of possible values of P. One way in which it is robust is if there may be mutations and laboratory errors, in which case P would have to be < 1. Taking P to equal 0.50 levies little penalty against a particular pair in which there is an apparent exclusion from direct parentage. Therefore taking P to be < 1 means that if the true parents are in the database then they will not be ruled out if there happen to be mutations and laboratory errors. And if the closest ancestors in the database are more remote than grandparents, they are likely to be identified because they will usually have the fewest mismatches of the lines considered.

When i and j are ancestors there are four possibilities: (1) The alleles of both inbreds i and j were passed to the hybrid. (2) inbred i came through but not inbred j. (3) inbred j came through but not inbred i, and (4) neither inbred came through. Assuming independence, these have respective probabilities P^2, $P(1 - P)$, $P(1 - P)$, $(1 - P)^2$. In the case $P = 0.50$, all of these probabilities equal 0.25.

An instance of the law of total probability (Sec. 5.3, Berry 1996) is that the probability of observing a hybrid's alleles is the average of the conditional probability of this event given the above four cases. The simplest of the four cases is the first possibility: Assuming the hybrid's alleles are passed down directly from both inbreds, the probability of observing the hybrid's genotype is either 1 or 0 depending on whether the hybrid shares both inbreds' alleles. (It is especially easy when both inbreds are homozygous.) The other three cases require an assumption regarding the possibility that an inbred's allele is not passed to the hybrid but is interrupted by a mutation, a laboratory error, or intervening breeding. We regard such an allele as being selected from all known alleles with probability 1/(number of alleles), where the number of alleles is the total number of alleles known to exist at the locus in question. An alternative approach would be to use the allelic proportions that are present in the database (or in another database). However, the lines in the database may not be randomly selected from any population. For example, a line that has been highly used in breeding would have many derivative lines in the database, in which case the frequencies of its alleles will be artificially inflated. Assuming equal probabilities for the various alleles at a given locus is robust in the sense that it is not affected by adding and dropping lines from the database.

There are many cases to consider when computing the probability of observing a hybrid's alleles, depending on the zygosity of the hybrid and the inbreds, and allowing for the possibility of missing alleles or "extra alleles" in the assessment of the hybrid and inbred genotypes. These possibilities are too numerous to list. Instead we give three simple examples. All the examples have homozygous inbreds, the most common case. And each of the three hybrids has two alleles, again the most common case. We suppose that the measured alleles for three SSRs and a particular trio of hybrid and ancestor inbreds are as we have indicated in Table 1.

For SSR 1 there are three known alleles, one in addition to alleles a and b that are listed for the three lines (hybrid, inbred i, and inbred j) in Table 1. For SSR 2 and SSR 3 there are two known alleles in addition to those listed. The calculations in the right half of Table 1 will now be explained. Implicit in calculating $P(\text{SSR}|i, j)$ is the assumption—required in both the numerator and denominator of Bayes' rule—that inbreds i and j are ancestors of the hybrid. Consider SSR 1. In case 1 above, both ancestors' alleles (as measured by the laboratory process) are assumed to pass to the index hybrid, and so in this case the hybrid is necessarily ab. The probability of observing the actual hybrid's genotype is 1 for case 1, as shown in Table 1. In case 2, we assume that inbred i's allele passes to the hybrid but inbred j's does not. Indeed, the hybrid has an a allele. The probability of observing a b as the other allele is 1/(number of alleles) = 1/3, as shown in Table 1. Case 3 is similar. In case 4, neither ancestor allele is passed to the hybrid; the probability of observing the hybrid's genotype (or any heterozygous genotype) is $2(1/3)(1/3) = 2/9$. Since $P = 0.50$, the overall (unconditional) probability in the rightmost column (17/36) is the simple average of the four cases, as indicated in Table 1.

For SSR 2 and SSR 3 the calculations are similar. For SSR 2 there is some evidence against pair (i, j) being ancestors,

TABLE 1

Probability of observing a hybrid's alleles using three sample SSRs and four possible combinations (cases) of alleles passed, assuming that inbreds *i* and *j* are ancestors of the hybrid

SSR	No. of alleles	Hybrid	Inbred *i*	Inbred *j*	Probability of observing the hybrid's genotype: Case 1 *i, j*	Case 2 *i, not j*	Case 3 *not i, j*	Case 4 *not i, not j*	Overall probability $P(SSR\|i, j)$
1	3	*ab*	*aa*	*Bb*	1	1/3	1/3	2/9	17/36
2	5	*bd*	*bb*	*Cc*	0	1/5	0	2/25	7/100
3	6	*ab*	*cc*	*Dd*	0	0	0	2/36	2/144

SSR, simple sequence repeat marker profile.

but it is not conclusive. For SSR 3 there is even less evidence favoring pair (i, j). It would not take many SSRs with evidence similar to that for SSR 3 to essentially rule out this pair—provided that other pairs are not similarly inconsistent.

To find the overall $P(SSRs|i, j)$, multiply the individual $P(SSR|i, j)$ over the various SSRs. There are purely computational issues to address. Each $P(SSR|i, j)$ is a number between 0 and 1. When there are a great many SSRs, the product of these numbers will be vanishingly small. To lessen problems with computational underflow, for each SSR we multiply $P(SSR|u, v)$ by the same constant for each pair (u, v): the inverse of the largest possible such probability. For example, since 17/36 is the largest probability for a heterozygous hybrid at an SSR having three alleles (as is the case for SSR 1 in Table 1), we multiply all factors $P(SSR|u, v)$ by 36/17. To eliminate remaining problems with underflow, we do calculations using logarithms (adding instead of multiplying) and take antilogs at the end.

The probability $P(SSR|u, v)$ is calculated for all (u, v) pairs and summed over all possible pairings in the database, including that for the inbred pair under consideration: (i, j). This gives the denominator in the expression for $P(i, j|SSRs)$.

To determine the probability that any particular inbred, say inbred i, is the closest ancestor of the index hybrid, sum $P(SSR|i, v)$ over all inbreds v with $v \neq i$. Call this $P(i|SSRs)$. The maximum of $P(i|SSRs)$ for any inbred i is 1. But since there is one closest ancestor on each side of the family, the sum of $P(i|SSRs)$ over all inbreds i is 2. If there is a particular pair (i, j) for which $P(i, j|SSRs)$ is close to 1 then both $P(i|SSRs)$ and $P(j|SSRs)$ separately will be close to 1.

SSR data: DNA was extracted from 54 maize hybrids and from 386 maize inbreds. All of the hybrids and most inbreds are proprietary products of Pioneer Hi-Bred International; some important publicly bred inbred lines were also included. The inbred parents and grandparents of each hybrid were included within the set of inbreds. Other inbreds that were genotyped include many that are highly related by pedigree to parents and grandparents of the hybrids. The hybrids were chosen because each has a pedigree that is known to us and collectively they represent a broad array of diversity of maize germplasm that is currently grown in the United States ranging from early to late maturity.

A total of 195 SSR loci were used in this study following procedures described in Smith *et al.* (1997), but modified as described below. SSR loci were chosen on the basis that they individually have been shown to have a high power of discrimination among maize inbred lines and collectively they provide for a sampling of diversity for each chromosome arm. Of these SSR loci, the following numbers (in parentheses) were located on individual maize chromosomes as follows: 1 (35), 2 (26), 3 (22), 4 (20), 5 (16), 6 (9), 7 (6), 8 (18), 9 (12), and 10 (14); 17 SSR loci have not yet been mapped. The correlations among the loci are unknown and are irrelevant for our methodology.

Sequence data for primers that allow many of these (and other) SSR loci to be assayed are available at website http://www.agron.missouri.edu. All primers were designed to anneal and amplify under a single set of conditions for PCR in 10-μl reactions. Genomic DNA (10 ng) was amplified in 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-Cl (pH 8.3) using 0.3 units AmpliTaq Gold DNA polymerase (PE Corporation) oligonucleotide primer pairs (one primer of each pair was fluorescently labeled) at 0.17 μM and 0.2 mM dNTPs. This mixture was incubated at 95° for 10 min (hot start); amplified using 45 cycles of denaturation at 95° for 30 sec, annealing at 60° for 30 sec, extension at 72° for 85 sec; and then terminated at 72° for 10 min. A water bath thermocycler manufactured at Pioneer Hi-Bred International was used for PCR reactions. PCR products were prepared for electrophoresis by diluting 3 μl of each product to a total of 27 μl using a combination of PCR products generated from other loci for that same maize genotype (multiplexing) and/or dH20. Dilution of 1.5 μl of this mixture to 5 μl with gel loading dye was performed; it was then electrophoresed at 1700 V for 1.5 hr on an ABI model 377 automated DNA sequencer equipped with GENESCAN software v. 3.0 (PE-Applied Biosystems, Foster City, CA).

PCR products were sized automatically using the "local Southern" sizing algorithm (Elder and Southern 1987). After sizing of PCR products using GeneScan, alleles were assigned using Genotyper software (PE-Applied Biosystems). Generally, allele assignations for each locus were made on the basis of histogram plots consisting of 0.5-bp bins. Breaks between the histogram plots of >1 bp were generally considered to constitute separation between allele bins; however, other criteria, such as the presence of the nontemplate-directed addition of adenine (+A addition) and naturally occurring 1-bp alleles, were used on a marker-by-marker basis to define the allele dictionary. All allele scores were made without knowing the identities of the maize genotypes.

RESULTS

Table 2 presents the probability of closest ancestry of the top five ranking inbred lines for each of 5 hybrids at $P = 0.50$ (Table 2A) and $P = 0.99$ (Table 2B). Probabilities of ancestry are shown for all 54 hybrids and the top ranking inbreds in Figure 1: $P = 0.50$ (Figure 1a) and $P = 0.99$ (Figure 1b). Results for the hybrids presented in Table 2 are featured at the top of Figure 1.

TABLE 2

Probability of ancestry of five hybrids using data obtained from 50, 100, and 195 SSR loci

	50 loci			100 loci			195 loci		
Hybd.	Inbd.	Prob.	SE	Inbd.	Prob.	SE	Inbd.	Prob.	SE
				A. Assuming $P = 0.50$					
3417	SP1	0.9607	0.0125	P1	0.8719	0.0252	P1	1.0000	E-07
	P2	0.8077	0.1965	P2	0.8141	0.2235	P2	0.9957	0.0033
	D2P2	0.1016	0.1938	D1P2	0.1839	0.2235	D1P2	0.0043	0.0053
	D1P2	0.0907	0.0927	SP1	0.1218	0.025	D2P2	E-06	E-06
	P1	0.032	0.0125	D1P1	0.0009	0.0002	SP1	E-06	E-07
3525	P1	0.8545	E-07	P1	0.9999	<E-20	P1	1.0000	<E-20
	P2	0.8188	E-07	P2	0.5437	<E-20	P2	0.9635	0.0528
	D1P2	0.1699	E-07	D1P2	0.4563	<E-20	D1P2	0.0365	0.0528
	GP1	0.1441	E-07	GP1	E-07	E-18	SP1	E-15	<E-20
	GP2	0.0110	E-08	SP1	E-07	<E-20	GP2	E-16	<E-20
3556	P1	1.0000	E-06	P1	0.9999	E-10	P1	1.0000	<E-20
	P2	0.9616	E-08	P2	0.9997	E-10	P2	1.0000	<E-20
	D1P2	0.0340	E-10	D1P2	0.0003	E-14	D1P2	E-09	<E-20
	GP2	0.0043	E-09	D2P2	E-05	E-15	D2P2	E-14	<E-20
	D2P2	0.0002	E-10	D3P2	E-06	E-17	GGP2	E-17	E-17
3905	D1P1	0.9822	E-08	D1P1	0.9803	0.0058	P1	1.0000	E-08
	SP2	0.4927	E-07	SP2	0.6280	0.0976	D1P2	1.0000	E-06
	D2P2	0.2536	E-07	D1P2	0.2321	0.0617	D2P2	E-06	E-06
	D1P2	0.1622	E-07	D2P2	0.1317	0.0372	P2	E-07	E-13
	P2	0.0565	E-07	P1	0.0197	0.0058	D3P2	E-10	E-16
3940	P2	0.9997	0.0001	P2	0.9999	E-05	P2	1.0000	E-09
	D1P2	0.9203	0.0009	P1	0.9970	0.0011	P1	1.0000	E-09
	P1	0.0648	E-05	D1P2	0.0030	0.0011	D1P2	E-11	E-11
	D1P1	0.0127	E-05	D2P2	0.0001	E-05	DP1P2	E-17	E-17
	DP1P2	0.0014	0.0009	DP1P2	0.0001	E-07	D2P2	E-19	E-18
				B. Assuming $P = 0.99$					
3417	SP1	0.9995	0.0001	P1	0.9999	E-05	P1	0.9999	E-08
	P2	0.8836	0.1658	P2	0.9938	0.0107	P2	0.9999	E-08
	D1P2	0.0722	0.1029	D1P2	0.0061	0.0107	D1P2	E-11	E-11
	D2P2	0.0441	0.0628	D1P1	E-05	E-06	D2P2	E-14	E-14
	P1	0.0004	0.0001	SP1	E-05	0	SP1	E-20	E-21
3525	P1	0.9999	0	P1	0.9999	0	P1	1.0000	0
	P2	0.8991	0	D1P2	0.9719	0	P2	0.6135	0.4446
	D1P2	0.1008	E-11	P2	0.025	0	D1P2	0.3864	0.4446
	GP1	E-05	0	D2P2	E-20	0	GP2	E-48	0
	GP2	E-06	E-17	SP1	E-24	0	D2P2	E-49	0
3556	P1	1.0000	0	P1	1.0000	0	P1	0.9999	0
	P2	0.9996	0	P2	0.9999	0	P2	0.9999	0
	D1P2	0.0003	0	D1P2	E-09	0	D1P2	E-22	0
	D1P1	E-11	0	D3P1	E-21	0	D2P1	E-49	0
	D2P1	E-13	0	D2P1	E-21	0	D3P1	E-54	0
3905	D1P1	0.9999	0	D1P1	0.9999	E-08	P1	1.0000	E-09
	P2	0.9992	0	P2	0.9999	E-06	P2	0.9947	E-09
	SP2	0.0006	0	D1P2	E-06	E-06	D1P2	0.0052	E-11
	D1P2	E-05	0	SP2	E-07	E-13	D2P2	E-18	E-18
	D2P2	E-06	0	D2P2	E-09	E-10	D1P1	E-25	E-25
3940	P2	0.9999	E-08	P2	1.0000	E-08	P1	1.0000	E-09
	D1P2	0.9999	E-08	P1	0.9999	E-05	P2	1.0000	E-09
	P1	E-06	E-13	D1P2	E-05	E-05	D1P2	E-24	E-24
	D1P1	E-08	E-15	D2P2	E-12	E-11	DP1P2	E-44	E-44
	DP1P2	E-12	E-12	DP1P2	E-21	E-21	D2P2	E-50	E-49

Hybd., hybrid; Inbd., inbred; Prob., probability; SE, standard error, referring to the variability in the results of the runs; P1, parent one; P2, parent two; SP1 SP2, full sibling of parent one parent two; D*i*P1/D*i*P2, derivatives of parent one parent two, index *i* for distinct inbred lines; DP1P2, derivatives of both parent one and parent two.

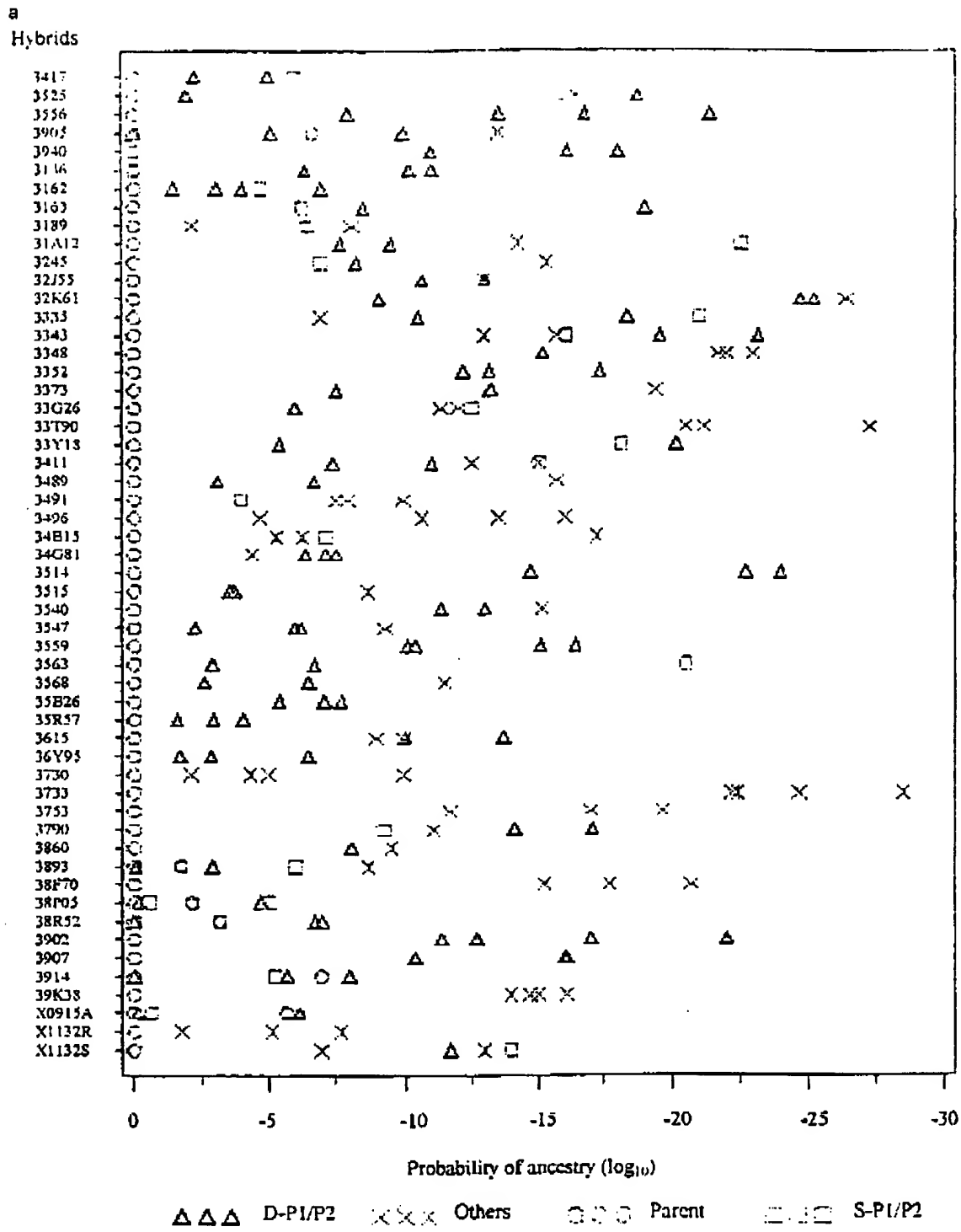


FIGURE 1.—(a) Probabilities of ancestry, assuming $P = 0.50$, for all 54 hybrids and top ranking inbreds—those with probability of ancestry at least 10^{-6}. (b) Probabilities of ancestry, assuming $P = 0.99$, for all 54 hybrids and top ranking inbreds—those with probability of ancestry at least 10^{-6}.

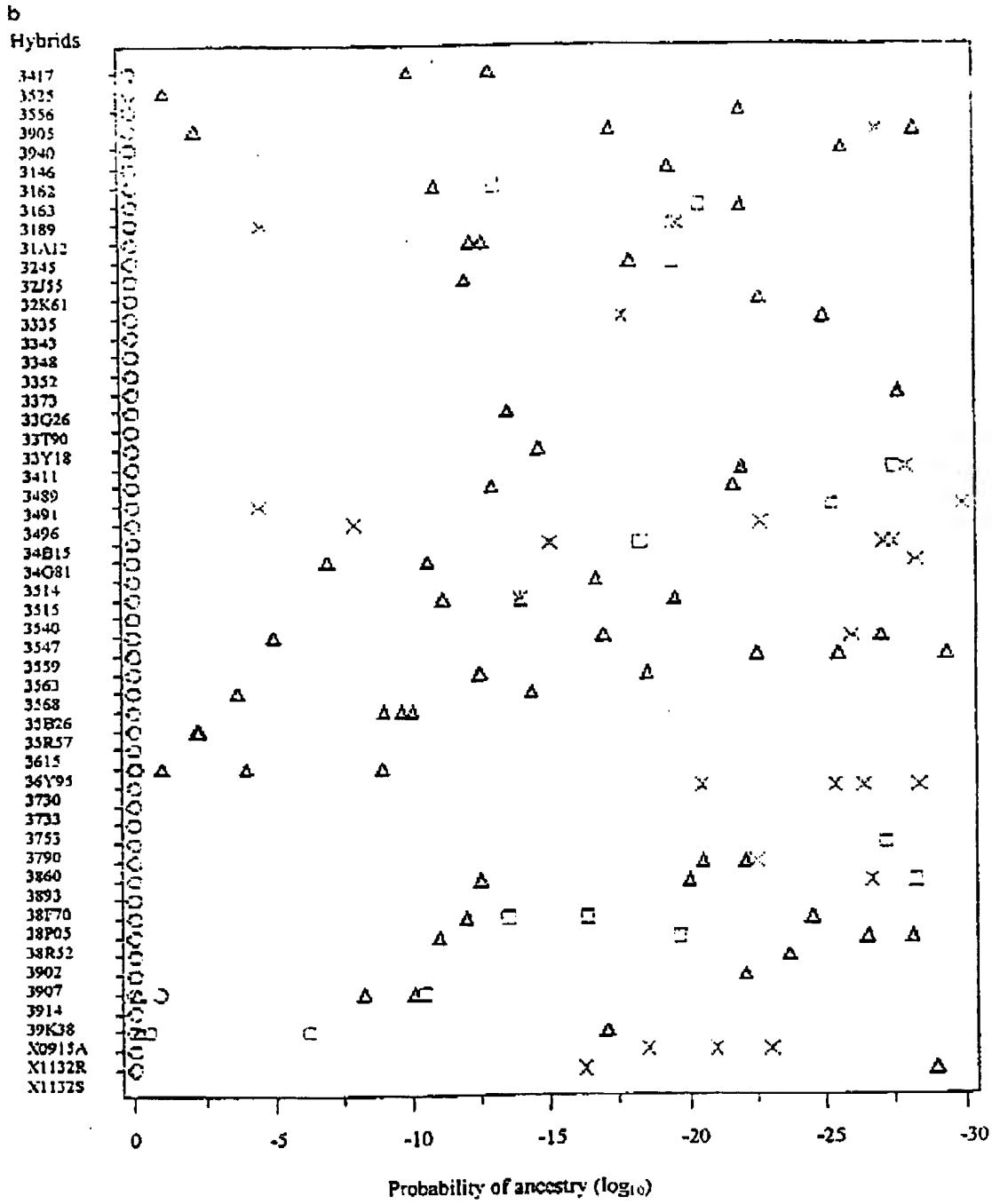


FIGURE 1.—*Continued.*

When the algorithm used $P = 0.50$, the two correct parents were identified as highest in probability for 48 (89%) hybrids (Figure 1). For each of 6 hybrids (3898, 38P05, 38R52, 3905, 3914, and X0915A), one parent ranked in the top two places. The other parent was supplanted either by a sister inbred or by an inbred that was a direct progeny of that parent. Overall, 102 (94%) of 108 parental inbreds were correctly identified. For hybrids where both parents ranked first or second, the range of probabilities for parental lines that ranked first from among all other inbreds ranged from 1.0000 to 0.9997; parental lines ranking second ranged from

1.0000 to 0.9653. For 35 hybrids, both parents had probabilities of ancestry in excess of 0.999. Probabilities of ancestry for nonparents that ranked in first or second places were from 0.9999 to 0.7054. For the majority of hybrids, the probability of the third and highest ranked nonparental inbred was at or below E-06. This indicates that there is usually very little uncertainty about closest ancestors.

When the algorithm used $P = 0.99$ to examine each of the 54 hybrids, both parents were correctly identified for 52 (96%) of hybrids and for 98% (102/104) of the parents across all hybrids (Figure 1). Two hybrids (3914 and X0915A), in which one parent was not ranked in the top two, were also in the subset not ranked in the top two assuming $P = 0.50$ (above). In both cases their ranks improved (both to third rank) and the actual parent was supplanted by an inbred that was a direct progeny of the corresponding parental line. For 49 hybrids, both parents had probabilities of ancestry in excess of 0.999. Among the 5 hybrids having a parent ranking second with a probability of ancestry below 0.999, the lowest of these probabilities was 0.8976 and the highest probability for a third ranking nonparent was 0.1023. For most hybrids the probability for the third and highest ranked nonparental inbred was at or below E-10.

Table 2 also addresses data analysis in circumstances where heterozygous loci occur in inbred lines or where a hybrid is scored for the presence of more than two alleles per locus. The presence of more than a single allele per locus in inbred lines is an infrequent occurrence in well-maintained inbred development and seed increase programs but is possible because ~3–5% of loci can still be segregating and unintended pollination from genotypes not designated as parents of the hybrid can occur. For hybrids, more than two alleles per locus can be scored when DNA is extracted from a bulk of individual plants and because inbred parents are not homozygous due either to residual heterozygosity or to contamination or because one or more direct parents of the hybrid are themselves hybrids. The presence of more than one allele per locus in an inbred line and more than two alleles per locus in a hybrid therefore can be accommodated by multiple runs of the algorithm, each with a random choice of two alleles per locus. Consequently, standard errors in the case of analyzing data from 195 loci tend to be very small because there were few loci where an inbred or hybrid sample (from a bulk of individual plants) was scored for more than two alleles.

Marshall *et al.* (1998) have drawn attention to errors that can be encountered in genotyping surveys. These errors include missing data, null alleles, and typing errors. We therefore investigated the robustness of the algorithm by examining the effects of modifications in the data for five hybrids (3417, 3525, 3556, 3905, and 3940). First, we reduced the number of SSRs used, from the full set of 195 to 100 and then to 50 (Table 2). Use of 50 loci generated incorrect rankings of one parent for each of two hybrids (3417 and 3940) and for both parents of one hybrid (3905). All of these most highly ranked nonparental inbreds were closely related to the true parents for each of the respective hybrids; six different inbred lines were involved. Four were direct progeny of the true parents (one with additional backcrosses from the true parent) and two were full sisters (from a cross of highly related inbreds) of the actual parent of the hybrid. Using 100 loci resulted in correct parental rankings for all hybrids except for 3905 where neither parent ranked in first or second place. Four inbreds outranked the true parents of 3905. All four nonparents were closely related to the respective true parents; three were direct progeny of the true parent of the hybrid (one with additional backcrossing to that parent) and one was a full sister of the true parent. Use of data from all 195 loci corrected the placement for one of the parents of hybrid 3905. Two inbreds that were not parents of this hybrid remained ranked more highly than one of the true parents. Both were direct progeny of that parent, and one of these inbreds had additional backcrossing to that parent in its pedigree.

To address the consequences of laboratory and other sources of error, we artificially compromised data quality beyond the level originally provided by eliminating specific proportions of alleles that had been scored (establishing scenarios where various numbers of SSR alleles were not scored) and by misscoring other alleles (establishing scenarios where various numbers of SSR alleles were scored incorrectly). We also combined the scenarios of missing data and wrongly scored data. Table 3 contains a summary of the results of making these modifications in the data. For all modifications we used data from all SSR loci and we also randomly chose SSR loci to create subsets of 50 and 100 loci. In each case, the program was run 20 times for each hybrid/set of loci. When all 195 loci were examined, replications differed only according to the particular choice of alleles for loci where more than two alleles had been scored.

To evaluate robustness in the face of missing data or mistyped data, we simulated individual and combined categories of these data in the hybrid and all inbred lines at levels of 2, 5, 10, and 25% of the alleles for each of five hybrids and all inbreds beyond the level of error as originally scored by the laboratory. We examined the effects of these levels and types of error for three sizes of database: 50 loci, 100 loci, and all 195 scored loci. The same five hybrids considered in Table 2 were investigated: 3417, 3525, 3556, 3905, and 3940. One of these hybrids (3905) was chosen because one of its parents did not rank among the top two places even when the complete and unmodified data from all SSR loci were used.

Examples of robustness in the face of additional error

TABLE 3

Number of parents ranked in first and second positions (maximum is 2)

Type of simulated data	% level change	No. of loci:	Hybrid 3417 50	3417 100	3417 195	3525 50	3525 100	3525 195	3556 50	3556 100	3556 195	3805 50	3805 100	3805 195	3940 50	3940 100	3940 195	Mean % max.
Missing	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	1	2	2	2	2	0	0	1	1	2	2	73
	5		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	10		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	25		0	2	2	1	1	1	2	2	2	0	0	1	1	1	1	57
	Mean % max.		40	100	100	90	80	90	100	100	100	0	0	50	50	90	90	
Misscored	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	5		1	2	2	2	1	2	2	2	2	0	0	1	1	2	2	73
	10		1	1	2	2	2	2	2	2	2	0	0	1	1	2	2	73
	25		1	0	2	1	1	2	1	2	2	0	1	1	1	2	2	63
	Mean % max.		50	70	100	90	80	100	90	100	100	0	10	50	50	100	100	
Missing plus misscored	0		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	2		1	2	2	2	2	2	2	2	2	0	0	1	1	2	2	77
	5		1	2	2	2	1	1	2	2	2	0	0	1	1	2	2	70
	10		1	2	2	1	2	2	2	2	2	0	1	1	1	2	2	7
	25		0	1	2	0	0	0	0	0	0	0	0	1	0	0	1	17
	Mean % max.		40	90	100	70	70	70	80	80	80	0	10	50	40	80	90	
Overall mean			48	87	100	83	77	87	90	93	93	0	7	50	47	90	93	

Hybrids considered are the same as those in Table 2.

for five hybrids using subsets of 50 and 100 loci and all loci are shown in Table 3 where numbers of parents ranking into the top two places are presented. Degradation in the preferential ranking of parent inbreds at a level of 25% additional missing data was shown for one hybrid (3525) with usage of 50, 100, or all SSR loci. Degradation in the preferential ranking of parent inbreds at a level of 25% additional misscored data was shown for hybrid 3556. When both additional levels of missing and misscored data were simulated, degradation in the ability to preferentially rank inbred parents occurred for all hybrids and for all sets of SSR (50, 100, and 195 loci) except for hybrid 3417 when data from 195 SSR loci were used. Over all five hybrids, use of 100 loci improved robustness from the use of 50 loci; use of 193 loci further improved robustness for four hybrids (3417, 3525, 3905, and 3940). The degree of improvement was small, except for hybrid 3905.

We also ranked inbreds according to their probability of ancestry of hybrids when both parents and all inbred derivatives and full-sister inbreds of the respective inbred parents for each hybrid were excluded from the analysis. The results are too voluminous to present here but can be summarized as follows: Using $P = 0.50$, a grandparent of each respective hybrid ranked into first place for 41 (76%) hybrids; probabilities ranged from 0.4976 to 1.0 and most were above 0.9999. Other classes of inbreds that ranked in first position for probability of ancestry were inbreds derived directly by pedigree from a grandparent of the respective hybrid (DGP) for 13% of hybrids, inbreds derived directly by pedigree from a great-grandparent of the respective hybrid (DGGP) for 9% of hybrids, and one class (2% of hybrids) with an inbred ranked into first place that was directly related by pedigree to the great-great-grandparent of that hybrid. Inbreds that ranked in second position were related to the respective parents of the hybrid as follows: Thirty-one (57% of hybrids) were a grandparent of the respective hybrid, 11 (20%) were classed as DGP, 7 (13%) were DGGP, 1 (2%) was class DGGGP, and 4 (7%) were a great-grandparent (GGP) of the respective hybrid. Over all hybrids, two of the four grandparents ranked into first and second positions for 23 (43% of hybrids); three grandparents ranked into the first three positions for 5 (9% of hybrids). There were no instances where all four grandparents ranked into the first four positions. Thirty hybrids had a grandparent ranked into first position using $P = 0.99$. The number of grandparents ranked into the top five positions was 93 (compared to 108 when $P = 0.50$). The number of grandparents ranking into the top two positions was 55 (compared to 71 when $P = 0.50$). The mean probability of a grandparent that ranked into the first two positions was 0.9288 (SD = 0.1454) when $P = 0.50$ and 0.9980 (SD = 0.01041) when $P = 0.99$.

DISCUSSION

The prevalent use of paternity indices demonstrates that it is advantageous to have explicit probabilities of ancestry to distinguish among different pedigrees. Molecular marker profiles are rapidly becoming more extensive and cost effective to generate. Features that would advance the statistical analysis of molecular marker data to provide explicit probabilities of ancestry include the ability to calculate probabilities of ancestry where there is no *a priori* information as to the identity of one (usually the maternal) parent and robustness in the face of laboratory error.

Maize inbred lines and hybrids provide a very exacting set of materials for evaluating the discriminatory abilities of molecular data and statistical procedures that are employed to interpret those data. Hundreds of maize inbred lines of known pedigree together encompass a great diversity and complexity of pedigree relationships. Some inbred lines can be very highly related and genetically similar due to their derivation from common parentage including from parents that are themselves highly related. Consequently, relationship categories such as "sister" or "parent" when applied to maize inbreds usually refer to closer degrees of pedigree relationship and, thus, of germplasm and molecular marker profile similarity than those of the equivalently named classes of relationship for animal species. Most maize hybrids that are widely used in the United States today are constructed from pairs of inbred lines that are unrelated by pedigree, each inbred parent having been bred from a separate "pool" of germplasm. Various degrees of relatedness are possible between hybrids according to the pedigree relationships among their constituent inbred parents.

Using $P = 0.99$ in the algorithm is more specific for identifying parents than using $P = 0.50$. However, $P = 0.99$ is less robust for identifying other relatives, such as grandparents. When the algorithm was run at $P = 0.50$ there were 6 hybrids for which one parent did not rank among the top two most probable genotypes. For the remaining 48 hybrids the correct parents were identified even in circumstances where other candidate inbreds included not only full-sister lines bred from related parents but also inbreds even more closely related to the true parent by virtue of being backcross conversions of the inbred parent of the hybrid. For each of the 6 hybrids where a nonparent ranked above a true parent, that higher ranked inbred was always either a sister or progeny of the outranked true parent. The range of pedigree relationships as expressed by the Malécot coefficient of relatedness (Malécot 1948) that was encompassed by pairs of true parents and more highly ranked inbred relatives of the true parents was from 0.8390 to 0.9680. A coefficient of 0.8390 approximates a relationship between inbred A and A′ where

inbred A' has been bred from a cross of inbreds A and B with between one and two additional backcrosses of the parental inbred A. A Malécot coefficient of relationship of 0.9680 closely approximates a relationship between inbreds A and A'' where four additional backcrosses of parental inbred A follow the initial cross of inbreds A and B.

Running the algorithm at $P = 0.99$ in comparison to $P = 0.50$ raises the probability of ancestry for the parents while diminishing the probabilities for the third and lower ranking candidate inbred lines. Use of the algorithm at $P = 0.99$ increased both the percentage of hybrids with both parents ranked in the first two positions (from 89 to 96%) and the percentage of parental inbreds that were ranked first and second (from 94 to 98%). Two hybrids (3914 and X0913A) did not have both parents ranked first and second when the algorithm was run at $P = 0.99$. For both of these hybrids the nonparental inbred that outranked the true parent was itself a product by pedigree from the true parent that had been created by an additional four backcrosses of that parent; the Malécot coefficient of relationship between the parent of the hybrid and the inbred that outranked that parent for these two hybrids was 0.9636.

Robustness was tested by evaluating the effects of using data from different numbers of loci and by simulating additional levels of missing and misscored data up to combined levels of 25% error beyond that which was provided by the laboratory. From our experience, error rates of 5 to 10% can occur in SSR profiling of maize due chiefly to the combined effects of residual heterozygosity among seed lots and by deficiencies in the scoring of heterozygotes in hybrids. The additional levels of simulated error, therefore, include values (up to ~35% total error) that are well outside of our experience. For five hybrids that were examined, increasing the number of loci from 50 to 100 (with no additional missing or misscored data) did reduce the number of instances where inbreds that were not parents of a hybrid outranked the true parent from four to one. Nonetheless, all of these more highly ranked inbreds, although they were not themselves the true parents of the respective hybrid, were either direct progeny or full sisters of the true parent (Table 2). Consequently, if such degrees of error can be tolerated in respect of pedigrees for inbreds that are identified as parents of hybrids, then SSR data from 50 loci of equivalent discrimination ability are sufficient. Use of data from 50 loci also evidenced robustness in the face of up to 10% additional levels of either missing or misscored data; no degradation in the ability to identify a parent was apparent up to the level of 10% additional error except for 10% additional missing and misscored alleles for one hybrid (3525; Table 3). However, use of 100 loci increased the proportion of true parents that were correctly identified from 53% (for 50 loci) to 71% (mean correct parents over all levels of error; Table 3). Use of data from 195 loci provided greater resiliency against additional levels of error. However, use of data from 195 loci was unable to provide resiliency against the negative effects of adding combined levels (at 25%) of both missing and misscored data (Table 3). At the 25% level of additional poor data integrity, inbreds that were not related to the true parent of the hybrid outranked the true parent for four of the five hybrids. Levels of missing or misscored data should, therefore, be kept below 15–20% (assuming a level of 5–10% error in the data we analyzed prior to simulating additional error).

We have previously examined the pedigrees of inbreds that are ranked into the first two positions when the true parents are removed from the list of candidate inbred lines. Usually, direct progeny or full sisters of the true parents then rank most highly (data not presented). We therefore examined the rankings of inbreds with respect to their ranking and probability of inclusion in the ancestry of each hybrid after the removal, not only of the true parents, but also of the progeny of the true parents and any full sisters of the true parents. In these circumstances the grandparents of the hybrids are ranked predominantly into top positions. Using $P = 0.50$, a grandparent ranked into first position for 76% hybrids and into second position for 57% hybrids; with $P = 0.99$ a grandparent ranked into first place in 56% of hybrids. At $P = 0.50$ two grandparents ranked into first and second positions for 43% hybrids and into the first three positions for an additional 9% hybrids. Most of the remaining inbreds that ranked into the top two positions were progeny of the grandparent. A total of 108 grandparents ranked into the top five positions when $P = 0.50$; 93 ranked into these positions when $P = 0.99$. Seventy-one grandparents ranked into the top two positions when $P = 0.50$; 55 grandparents ranked into these positions when $P = 0.99$. The mean probability of a grandparent in the top two positions was 0.9288 (SD 0.1454) when $P = 0.50$ and 0.9980 (SD 0.0104) when $P = 0.99$. Our algorithm was written to identify pairs of ancestors; alternative algorithms could be tailored to identify all grandparents once parents had been identified and removed from the list of candidate inbreds.

We have demonstrated the capability and robustness of an algorithm that can be used to show probability of parentage in circumstances where the *a priori* pedigree identity of neither parent is known. Exclusions are taken into account, thereby allowing parentage to be shown even when the two parents are not represented in the database of molecular profiles that are examined. Heterozygous candidate parents can be accommodated. The number of loci that is necessary to provide a reliable basis of determining pedigree is dependent upon the degree of relatedness among parents and nonparents and upon the discriminatory ability of the marker system

in the species of interest. Using $P = 0.99$ compared to $P = 0.50$ preferentially identified more true parents and with a greater difference of probability to third placed nonparents. If there is reasonable assurance that the parents are among the candidate list of inbreds, then $P = 0.99$ should be used; if greater robustness is required, then $P = 0.50$ should be used.

Applications of our algorithm include the identification of pedigrees among individuals of plant or animal species where molecular profile datasets exist that can be interpreted in terms of segregating alleles at individual marker loci and that provide a sufficient power of discrimination. Capabilities to generate large datasets of suitable molecular profile data are already available and are increasing rapidly with the advent of single nucleotide polymorphisms. One further application of our algorithm is to assist in the protection of intellectual property that is obtained on plant varieties or upon specific dams or sires of animals through the determination of pedigrees.

LITERATURE CITED

Alderson, G. W., H. L. Gibbs and S. G. Sealy, 1999 Parentage and kinship studies in an obligate brood parasitic bird, the brown-headed cowbird (*Molothrus ater*), using microsatellite DNA markers. J. Hered. 90: 182–190.

Bein, G., B. Driller, M. Schurmann, P. M. Schneider and H. Kirchner, 1998 Pseudo-exclusion from paternity due to maternal uniparental disomy 16. Int. J. Leg. Med. 111: 328–330.

Berry, D. A., 1991 Inferences using DNA profiling in forensic identification and paternity cases (with discussion). Stat. Sci. 6: 175–205.

Berry, D. A., 1996 *Statistics: A Bayesian Perspective*. Duxbury Press, Belmont, CA.

Berry, D. A., and S. Geisser, 1986 Inferences in cases of disputed paternity, pp. 353–382 in *Statistics and the Law*, edited by M. H. DeGroot, S. E. Fienberg and J. K. Kadane. Wiley Publishing, New York.

Bockel, B., P. Nurnberg and M. Krawczak, 1992 Likelihoods of multilocus DNA fingerprints in extended families. Am. J. Hum. Genet. 51: 554–561.

Bowers, J. E. and C. P. Meredith, 1997 The parentage of a classic wine grape, Cabernet Sauvignon. Nat. Genet. 16: 84–87.

Chakraborty, R., T. R. Meagher and P. E. Smouse, 1988 Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. Genetics 118: 527–536.

Chakraborty, R., L. Jin and Y. Zhong, 1994 Paternity evaluation in cases lacking a mother and nondetectable alleles. Int. J. Leg. Med. 107: 127–131.

Devlin, B., K. Roeder and N. C. Ellstrand, 1988 Fractional paternity assignment: theoretical development and comparison to other methods. Theor. Appl. Genet. 76: 369–380.

Elder, J. K., and E. M. Southern, 1987 Computer-aided analysis of one dimensional restriction fragment gels, pp. 165–172 in *Nucleic Acid and Protein Sequence Analysis—A Practical Approach*, edited by M. J. Bishop and C. J. Rawlings, IRL Press, Oxford.

Ellstrand, N. C., 1984 Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. Am. Nat. 123: 819–828.

Goetz, K., and G. Fiedler, 1998 Assignment of individuals to populations using microsatellites. J. Anim. Breed. Genet. 115: 73–81.

Gunn, P. R., K. Tieleman, P. Stapleton and D. B. Krakowski, 1997 DNA analysis in disputed parentage: the occurrence of two apparently false exclusions of paternity, both at short tandem repeat (STR) loci, in the one child. Electrophoresis 18: 1850–1852.

Hamrick, J. L., and A. Schnabel, 1985 Understanding the genetic structure of plant populations: some old problems and a new approach, pp. 30–70 in *Population Genetics in Forestry*, edited by H. R. Gregorius. Springer-Verlag, Heidelberg, Germany.

Helminen, P., V. Johnsson, C. Einholm and L. Peltonen, 1991 Proving paternity of children with deceased fathers. Hum. Genet. 87: 657–660.

Lang, J. W., R. K. Aggarwal, K. C. Majumdar and L. Singh, 1999 Individualization and estimation of relatedness in crocodilians by DNA fingerprinting with a Bkm-derived probe. Mol. Gen. Genet. 238(1–2): 49–58.

Malécot, G., 1948 *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris.

Marshall, T. C., J. Slate, L. E. B. Kruuk and J. M. Pemberton, 1998 Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639–655.

Meagher, T. R., 1986 Analysis of paternity within a natural population of *Chamaelirium luteum* (L). Identification of most-likely male parents. Am. Nat. 128: 199–215.

Meagher, T. R., and E. Thompson, 1986 The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theor. Popul. Biol. 29: 87–106.

Miller, P. S., 1975 Selective breeding programs for rare alleles: examples from the Przewalski's horse and California Condor pedigrees. Conserv. Biol. 9: 1262–1273.

Primmer, C. R., M. T. Koskinen and J. Piironen, 2000 The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. Proc. R. Soc. Lond. B Biol. Sci. 267: 1699–1704.

Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA 94: 9197–9201.

Sefc, K. M., H. Steinkellner, J. Glossl, S. Kampfer and F. Regner, 1998 Reconstruction of a grapevine pedigree by microsatellite analysis. Theor. Appl. Genet. 97: 227–231.

Smith, J. S. C., E. C. L. Chin, H. Shu, O. S. Smith, S. J. Wall *et al.*, 1997 An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor. Appl. Genet. 95: 163–173.

Smouse, P. E., and T. R. Meagher, 1994 Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) gray (Liliaceae). Genetics 136: 313–322.

Taylor, A. C., A. Horsup, C. N. Johnson, P. Sunnucks and B. Sherwin, 1997 Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat *Lasiorhinus krefftii*. Mol. Ecol. 6: 9–19.

Thompson, E., and T. R. Meagher, 1987 Parental and sib likelihoods in genealogy reconstruction. Biometrics 43: 585–600.

Vankan, D. M., and M. J. Faddy, 1999 Estimations of the efficacy and reliability of paternity assignments from DNA microsatellite analysis of multiple-sire matings. Anim. Genet. 30: 355–361.

Wetton, J. H., D. T. Parkin and R. E. Carter, 1992 The use of genetic markers for parentage analysis in *Passer domesticus* (house sparrows). Heredity 69: 248–254.

White, E., J. Hunter, C. Dubetz, R. Brost, A. Bratton *et al.*, 2000 Microsatellite markers for individual tree genotyping: application in forest crime prosecutions. J. Chem. Technol. Biotechnol. 75: 923–926.

Communicating editor: Z-B. Zeng

Attorney Docket No. 1235

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Delmar Brenner	Date:	March 11, 2003
Serial No.:	09/559,690	Group Art Unit:	1638
Filed:	April 27, 2000	Examiner:	David T. Fox
For:	"INBRED MAIZE LINE PH0R8"		

Assistant Commissioner for Patents
Washington, D.C. 20231

RULE 132 DECLARATION
OF
DR. DINAKAR BHATTRAMAKKI

Sir:

I, Dinakar Bhattramakki, Ph.D., do hereby declare and say as follows:

1. I am skilled in the art of the field of the invention. I have a Ph.D. in Plant Molecular Genetics from the University of Illinois at Urbana-Champaign. I have a Bachelor of Science degree in Agricultural Sciences from the University of Agricultural Sciences, Bangalore, India. Since 1997 I have been engaged in the analysis of molecular markers for plants. I have supervised the Molecular Marker Applications lab at Pioneer Hi-Bred International, Inc. from January 2002 until the present.

2. I am familiar with the methods used in the analysis of Simple Sequence Repeat, SSR, marker data for inbred PH0R8 conducted at Pioneer Hi-Bred International, Inc. The analysis of the SSR profile of inbred PH0R8 may be accomplished without any undue experimentation. The SSR profile for inbred PH0R8 is attached hereto.

3. Means of performing this genetic marker profile are well known in the art. SSRs are genetic markers based on polymorphisms in nucleotide sequences. The PCR™ detection of SSRs is accomplished by using two oligonucleotide primers flanking the polymorphic segment of DNA. Amplification is accomplished through

Appendix D

repeated cycles of heat denaturation of the DNA followed by annealing of the primers to their complementary sequences at low temperatures, and extension of the annealed primers with DNA polymerase.

4. Markers are scored following amplification and gel electrophoresis of the amplification products. Scoring of marker genotype is based on the size or weight of the amplified fragment. While variation in the primer used or in laboratory procedures can affect the reported marker score, relative values remain constant regardless of the specific primer or laboratory used.

5. Primers that may be used to identify the SSR markers reported herein are publicly available and may be found in the Maize DB on the World Wide Web at agron.missouri.edu/maps.html (sponsored by the University of Missouri), in Sharopova et al. (Plant Mol. Biol. 48(5-6):463-481) and/or in Lee et al (Plant Mol. Biol. 48(5-6); 453-461). Markers shown for PH0R8 are the publicly available markers in the sources listed above for which PH0R8 was tested and shown to be homozygous.

6. Map information is provided by bin number as reported in the Maize DB. The bin number digits to the left of decimal point typically represent the chromosome on which such marker is located, and the digits to the right of the decimal typically represent the location on such chromosome.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: 3/10/03 By: _____

Dinakar Bhattramakki

09/559,690

Public Name of Marker	bin #	PH0R8 base pairs
phi427913	1.01	129.38
bnlg1014	1.01	125.88
phi056	1.01	255.30
bnlg1083	1.02	203.72
bnlg1127	1.02	96.21
bnlg1953	1.02	205.80
bnlg1429	1.02	184.58
bnlg1627	1.02	206.48
bnlg439	1.03	232.00
phi109275	1.03	126.31
bnlg1203	1.03	306.51
bnlg1484	1.03	145.92
bnlg2086	1.04	217.79
bnlg1886	1.05	145.64
bnlg1057	1.06	274.00
bnlg1041	1.06	192.27
bnlg1615	1.06	211.96
phi335539	1.08	88.58
phi423298	1.08	133.68
phi002	1.08	73.53
bnlg1331	1.09	121.03
phi011	1.09	226.98
phi308707	1.10	131.11
phi227562	1.11	322.45
phi064	1.11	94.65
phi402893	2.00	218.34
bnlg1017	2.02	195.57
bnlg2277	2.02	292.29
bnlg1064	2.03	200.92
bnlg1018	2.04	134.33
bnlg1909	2.05	304.05
bnlg1138	2.06	222.14
bnlg1396	2.06	136.79
bnlg1831	2.06	193.90
phi328189	2.08	121.35
phi427434	2.08	121.46
phi435417	2.08	214.27
phi127	2.08	123.97
bnlg1520	2.09	294.41
phi101049	2.10	238.19
phi453121	3.00	217.80
phi104127	3.01	169.96

09/559,690

Public Name of Marker	bin #	PH0R8 base pairs
phi404206	3.01	300.24
phi374118	3.02	225.95
bnlg1144	3.02	159.57
bnlg1647	3.02	154.26
bnlg1523	3.03	257.05
bnlg1019	3.04	179.98
bnlg1113	3.04	134.41
bnlg1452	3.04	84.44
bnlg1035	3.05	118.64
phi053	3.05	191.64
phi102228	3.06	126.84
bnlg1160	3.06	220.49
bnlg1951	3.06	121.12
bnlg2241	3.06	142.63
phi072	4.00	139.43
phi213984	4.01	303.16
phi295450	4.01	197.12
phi308090	4.04	216.06
phi096	4.04	234.58
phi438301	4.05	211.81
bnlg1159	4.05	141.65
bnlg1937	4.05	227.36
bnlg1265	4.05	225.57
phi079	4.05	177.74
bnlg1006	5.00	229.38
phi396160	5.02	298.12
phi109188	5.03	165.50
bnlg653	5.04	151.67
phi330507	5.04	140.46
phi331888	5.04	130.72
bnlg1208	5.04	118.88
bnlg1892	5.04	149.85
phi333597	5.05	216.51
phi085	5.06	250.16
bnlg1118	5.07	82.56
bnlg1711	5.07	176.65
phi423796	6.01	128.39
phi445613	6.05	96.95
phi299852	6.07	117.30
bnlg1740	6.07	230.27
bnlg1759	6.07	125.81

09/559,690

Public Name of Marker	bin #	PH0R8 base pairs
phi070; umc1063	6.07	83.47
phi034	7.02	140.96
bnlg2271	7.03	218.19
phi328175	7.04	128.31
phi260485	7.05	285.39
phi069	7.05	204.40
phi116	7.06	158.89
phi420701	8.00	297.77
bnlg1194	8.02	140.77
phi100175	8.03	145.34
bnlg2082	8.03	152.96
phi115	8.03	302.63
phi121	8.03	97.90
bnlg2046	8.04	306.22
bnlg1152	8.06	148.42
bnlg1065	8.07	215.69
phi015	8.08	79.78
phi233376	8.09	147.93
bnlg2122	9.01	237.09
phi032	9.04	236.63
phi108411	9.05	126.04
phi236654	9.05	117.08
phi041	10.00	214.97
phi96342	10.02	240.18
phi059	10.02	143.60
bnlg1079	10.03	176.05
bnlg1655	10.03	126.51
phi050	10.03	83.27
phi301654	10.04	128.33
phi062	10.04	161.07
phi323152	10.05	144.53
bnlg1074	10.05	183.85
bnlg1185	10.07	183.31
bnlg1450	10.07	210.36
phi109642	2.03/2.04	148.23
bnlg1720	1.09/1.10	236.44
phi448880	9.06/9.07	182.99

What is an "Essentially Derived Variety"?

The concept of essentially derived variety was introduced into the 1991 Act of the UPOV Convention in order to avoid plagiarism through mutation, multiple back-crossing and to fill the gap between Plant Breeder's Rights and patents, gap which was becoming important due to the development of the use of patented genetic traits in genetic engineering.

An essentially derived variety is a variety which is distinct and predominantly derived from a protected initial variety, while retaining the essential characteristics of that initial variety.

As indicated as an example in the UPOV Convention, essentially derived varieties may be obtained by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, back-crossing, or transformation by genetic engineering.

The commercialization of an essentially derived variety needs the authorization of the owner of the rights vested in the initial variety.

The concept of essentially derived variety does not at all abolish the Breeder's Exemption, as free access to protected plant varieties for breeding purposes is maintained. It is not a threat to biodiversity. On the contrary, it favors biodiversity, encouraging breeders developing and marketing original varieties.

Appendix E

UPOV Publication No. 644(E), Section 1

INTERNATIONAL CONVENTION

FOR THE

PROTECTION OF NEW VARIETIES OF PLANTS

of December 2, 1961, as revised
at Geneva on November 10, 1972,
on October 23, 1978, and
on March 19, 1991

adopted by the Diplomatic Conference
on March 19, 1991

reproduced from UPOV Publication No. 438(E)
issue No. 63 of "Plant Variety Protection"

1991 Act of the Convention

Article 12
Examination of the Application

Any decision to grant a breeder's right shall require an examination for compliance with the conditions under Articles 5 to 9. In the course of the examination, the authority may grow the variety or carry out other necessary tests, cause the growing of the variety or the carrying out of other necessary tests, or take into account the results of growing tests or other trials which have already been carried out. For the purposes of examination, the authority may require the breeder to furnish all the necessary information, documents or material.

Article 13
Provisional Protection

Each Contracting Party shall provide measures designed to safeguard the interests of the breeder during the period between the filing or the publication of the application for the grant of a breeder's right and the grant of that right. Such measures shall have the effect that the holder of a breeder's right shall at least be entitled to equitable remuneration from any person who, during the said period, has carried out acts which, once the right is granted, require the breeder's authorization as provided in Article 14. A Contracting Party may provide that the said measures shall only take effect in relation to persons whom the breeder has notified of the filing of the application.

CHAPTER V
THE RIGHTS OF THE BREEDER

Article 14
Scope of the Breeder's Right

(1) [Acts in respect of the propagating material] (a) Subject to Articles 15 and 16, the following acts in respect of the propagating material of the protected variety shall require the authorization of the breeder:

(i) production or reproduction (multiplication),

(ii) conditioning for the purpose of propagation,

(iii) offering for sale,

(iv) selling or other marketing,

(v) exporting,

(vi) importing,

(vii) stocking for any of the purposes mentioned in (i) to (vi), above.

(b) The breeder may make his authorization subject to conditions and limitations.

(2) [Acts in respect of the harvested material] Subject to Articles 15 and 16, the acts referred to in items (i) to (vii) of paragraph (1)(a) in respect of harvested material, including entire plants and parts of plants, obtained through the unauthorized use of propagating material of the protected variety shall require the authorization of the breeder, unless the breeder has had reasonable opportunity to exercise his right in relation to the said propagating material.

(3) [Acts in respect of certain products] Each Contracting Party may provide that, subject to Articles 15 and 16, the acts referred to in items (i) to (vii) of paragraph (1)(a) in respect of products made directly from harvested material of the protected variety falling within the provisions of paragraph (2) through the unauthorized use of the said harvested material shall require the authorization of the breeder, unless the breeder has had reasonable opportunity to exercise his right in relation to the said harvested material.

(4) [Possible additional acts] Each Contracting Party may provide that, subject to Articles 15 and 16, acts other than those referred to in items (i) to (vii) of paragraph (1)(a) shall also require the authorization of the breeder.

(5) [Essentially derived and certain other varieties] (a) The provisions of paragraphs (1) to (4) shall also apply in relation to

(i) varieties which are essentially derived from the protected variety, where the protected variety is not itself an essentially derived variety,

(ii) varieties which are not clearly distinguishable in accordance with Article 7 from the protected variety and

(iii) varieties whose production requires the repeated use of the protected variety.

(b) For the purposes of subparagraph (a)(i), a variety shall be deemed to be essentially derived from another variety ("the initial variety") when

(i) it is predominantly derived from the initial variety, or from a variety that is itself predominantly derived from the initial variety, while retaining the expression of the essential characteristics that result from the genotype or combination of genotypes of the initial variety,

(ii) it is clearly distinguishable from the initial variety and

(iii) except for the differences which result from the act of derivation, it conforms to the initial variety in the expression of the essential characteristics that result from the genotype or combination of genotypes of the initial variety.

(c) Essentially derived varieties may be obtained for example by the selection of a natural or induced mutant, or of a somaclonal variant, the selection of a variant individual from plants of the initial variety, backcrossing, or transformation by genetic engineering.

Article 15
Exceptions to the Breeder's Right

(1) [Compulsory exceptions] The breeder's right shall not extend to

(i) acts done privately and for non-commercial purposes,

(ii) acts done for experimental purposes and

(iii) acts done for the purpose of breeding other varieties, and, except where the provisions of Article 14(5) apply, acts referred to in Article 14(1) to (4) in respect of such other varieties.

(2) [Optional exception] Notwithstanding Article 14, each Contracting Party may, within reasonable limits and subject to the safeguarding of the legitimate interests of the breeder, restrict the breeder's right in relation to any variety in order to permit farmers to use for propagating purposes, on their own holdings, the product of the harvest which they have obtained by planting,